

US nuclear elephants stay white

The continuing dispute in the United States about licences for two completed nuclear plants points to a failure to reconcile local and national interests — and to the need for federal leadership.

JUST before the US Congress adjourned for its summer break last week, it nearly (but not quite) found time to finish off the nuclear power industry in the United States. The immediate issue is whether two long-since completed nuclear plants, Shoreham on Long Island and Seabrook, across the border with New York in New Hampshire, should be given operating licences by the Nuclear Regulatory Commission (NRC), the federal agency with responsibility for such questions. Apparently on the principle that even if you have a dog, you still need to bark yourself, the House of Representatives solemnly considered whether NRC should be allowed to decide what conditions must be satisfied by those seeking licences for nuclear plants. Luckily, it decided (by 256 votes to 150) to let NRC do its job, but because the issue arose as an amendment to another bill, there is every likelihood that it will come up again.

What is at stake, in the United States and wherever local communities are asked to make sacrifices for the greater good, is the balance to be struck between local and national interests. Shoreham and Seabrook show that the United States still has to find a sensible way of arriving at wise conclusions.

The circumstances are curious. After the accident at Three Mile Island, NRC decided in 1980 that it would be a condition of future operating licences that there should not merely be a plan for evacuating local people in case of accidents (always a condition), but also that local authorities should have been convinced that the plans made sense. The principle is splendidly democratic. What better assurance can there be of local people's contentment than their representatives' acknowledgement that even emergencies can be dealt with safely? The snag is that local people and authorities have nothing much to gain by giving their consent. Local politicians are especially in a cleft stick; saying yes will alienate voters who believe nuclear power to be an abomination, but the benefits (electricity from plants that must otherwise stand idle at the public expense) are neither easily recognizable nor identifiable with themselves.

The evacuation plans at both Shoreham and Seabrook have been declared inadequate by at least some of the local authorities concerned. So NRC now plans to change the rules to dispense with local agreement on evacuation plans. Last week's amendment, if accepted, would have put a stop to that.

Minuet

Now will follow a predictable minuet. NRC will no doubt take back onto its own shoulders responsibility for deciding whether operators' evacuation plans are adequate, will be accused of moving the goalposts in the continuing dispute about the siting of nuclear plants in the United States and its decision will be challenged in the courts. Those who want the Shoreham and Seabrook plants to remain white elephants have powerful allies, including the governors of Massachusetts (Mr Michael S. Dukakis) and of New York (Mr Mario Cuomo).

This means that at least the Seabrook plant will be a vivid political issue next February, when Dukakis tries to persuade the voters of New Hampshire to support his bid to become the Democratic candidate in next year's presidential election. Piquantly, New Hampshire wants Seabrook to be licensed.

Shoreham, completed four years ago, is an even more complicated case, partly because the local population density is greater and partly because Cuomo remains many Democrats' favourite candidate for 1988 even though he has said he will not put up for election.

In the circumstances, the owners of the two plants had better reconcile themselves to further postponement. Ironically, electricity consumers will not have to foot the bill for the extra interest that will have to be paid on the best part of \$10,000 million; the rules require the electricity utilities to add the costs of keeping the plants idle to the construction cost, paying interest on what they owe their creditors with newly borrowed capital. Nobody seems to care that this process of rolling up accumulated debt will eventually entail either higher electricity prices or the bankruptcy of the utilities.

Rotten way

This is a rotten way to decide the important social issue of how best to share the disadvantages of economic development among those who are also its beneficiaries. NRC was mistaken, in 1980, to give local interests such a powerful voice, especially when it is plain that the consequences of a nuclear accident serious enough to warrant evacuation must be a federal responsibility. NRC should have guessed that the result would be to undermine its own regulatory authority, nicely illustrated last week by the decision of the state of Massachusetts to dispute the new evacuation plan filed by the operators of the Pilgrim reactor at Plymouth, Massachusetts, on the grounds that the plan assumes the availability of buses owned by a private transportation company.

NRC is right to want to rectify its mistake. But it would have a better chance of success if it more openly explained its difficulties; if it were better placed to demonstrate that the US nuclear industry has rid itself of laxities — such as that last year, when the technicians operating the Peach Bottom reactor in Illinois were found to be using their control room as a cross between a recreational facility and a dormitory (shades of Chernobyl); and if it had a better tale to tell about the economics of nuclear power in the United States. The underlying issue, however, is not that of whether nuclear energy now makes technical and economic sense, but that of reconciling local and national interests.

Novel projects as different as highways, oil refineries and nuclear power stations, economic necessities though they may be, are undoubted nuisances for those who live near them. On the face of things, it may seem equitable that those living near a completed nuclear power station should have the right to decide whether it should be allowed to operate, but that is a superficial judgement that would give local objectors powers they do not enjoy in connection with other kinds of schemes. In reality, both Shoreham and Seabrook have a regional importance beyond the disputed evacuation zones. It is difficult to believe that the objections to both schemes would have been as widely supported in New York and Massachusetts if electricity consumers were already paying the full costs of keeping workable nuclear plants in mothballs. And it is far from equitable that the two states concerned should be willing to risk the bankruptcy



of the utilities concerned, when the costs would fall not on the local people but on the bondholders.

But local people are entitled to expect that NRC has done everything it can to ensure that the plants will operate safely, and that it has workable plans for dealing with emergencies. If NRC takes back responsibility for deciding these questions, it could do worse than carry the argument to the people who now think themselves to be in serious danger. In the present climate, that will be a difficult undertaking. Ideally, the administration of which NRC is a small part should do what it can to help. But with an election coming up, that is a lot to ask for. □

UNCTAD's good news

Rich and poor countries seem nearer to an understanding on mutual assistance, but will it stick?

THE United Nations Commission on Trade and Development (UNCTAD), which meets in full session every four years, seems to have had a productive session last month in Geneva. For an organization best known as a talking shop, and an acrimonious one, that comes as a surprise.

Why, this year, should the climate apparently have changed? There have been two important developments. The rich countries, having mostly failed to provide aid on the scale for which the poor have been asking, have found that the poor have been able to find some of the cash they needed by borrowing from the West's commercial banks and then letting the loans turn sour. The fact that the funds have been taken not from public revenues but from the pockets of the banks' shareholders is irrelevant. What is plain is that this process will not easily be repeated, that neither financial aid nor commercial bank loans are likely to be distributed among the poor countries in the way that makes most economic sense and that the best long-term solution, for rich and poor alike, is to enable the poor countries to trade their way out of poverty. That is what the developed world seems to have been saying at Geneva.

At the same time, the poor countries seem to have gone a long way to accepting the Western idiom in which capital and interest are different kinds of money sums, in which political autonomy has its roots in solvency and in which the provision of basic skills, in agriculture and by education, is a necessary step in the development of a stable society.

What does UNCTAD's new amity imply for the years ahead? The most dramatic of the proposals canvassed at Geneva, by the United States, is that there should be an international agreement to abolish agricultural subsidies (see *Nature* 328, 188; 1987) over a period of ten years. Part of the objective is to give poor countries a better chance to make their way in the world by selling what they most easily produce, although the United States appears also to have calculated that getting rid of subsidies would be the most effective way of giving US farmers export markets in protected Europe and Japan.

Whatever the US motives, there can be no doubt that the poor countries would also benefit. But there is only the slimmest prospect that the rich countries will agree among themselves to follow such a course. Nor will they easily be persuaded to abolish the "multi-fibre agreement" used to protect the textile industries of the West from cheaper imports from elsewhere.

That is why UNCTAD's one successful meeting should not mislead. It will be splendid if the meeting can be taken as a sign that the rich countries of the world are coming round to the notion that the free trade they espouse (with limitations and reservations) among themselves is logically indivisible, and that the poor countries should be helped to share its benefits (and learn its disciplines). To be sure, some of the governments of poor countries are at present among the worst offenders. But is there hope of persuading them to change their ways while rich countries conspire to prevent them earning what they might by honest trade? □

Economy in space

Britain's National Space Centre should be put on the back burner for a few decades.

MR Roy Gibson, director of the British National Space Centre, did the decent thing when, last week, he quit his job. Gibson has spent nine months pleading with the government that employed him for a budget commensurate with the grand expectations of his post, magicked out of thin air early in 1986 by Mr (now Sir) Geoffrey Pattie, a somewhat oafish minister in the pre-election Thatcher government. Gibson did not get his way. To be fair to everybody, it was never clear whether the National Space Centre was meant as Britain's proof that it would make a mark in space (spending new money on the way) or as a relabelling exercise in which various activities, some military and some civil, would be lumped together and coordinated to make them more visible. Did Pattie win his senior colleagues' approval for his little scheme by telling them he meant the second while telling people like Gibson that he meant the first? The British will not know for 30 years when some public records are made public), but Pattie has been dropped from the government; it is tough that Gibson should have to resign his job as well.

The dilemma remains for Britain, and for many other countries like it, of what part it should play in space. Gibson's resignation has sent British newspapers scurrying through their files to remind themselves that, a quarter of a century ago, Britain was busily developing the Bluestreak military rocket as if it were almost one of the big boys. Why, the British press has been asking, was that technically successful project abandoned when the total cost was comparatively small, only to become the basis of the mostly successful and titularly European (but largely French) civilian rocket Ariane?

In reality, Bluestreak's demise had as much to do with the military calculation that the deployment of conspicuous immobile missiles would make Britain more and not less vulnerable as with the arrival of the first Wilson government in 1964 or the money that had been spent. But the rhetorical question the newspapers have been asking deserves a straight answer: Britain (and countries like it) does not have the technical resources, and in particular the technical manpower, to make projects like these succeed while keeping the economy ticking over. Preoccupied though Gibson may have been with other matters, even he will have remarked on last week's stark proof of that, the 1 per cent increase of British interest rates on Thursday and the precipitous decline of the London stock market that followed.

But what connection can there be between Britain's capacity to play a daring role in space and the evident continuing fragility of its economy? The lack of technical competence is the connecting thread. This week's economic statistics will show whether the Bank of England has taken fright because Britain's trade figures have gone awry again or because bank lending shows inflationary trends once more. Whatever the case, the explanation is the same. Britain is no longer a substantial source of technically advanced manufactured products so distinctive that international customers have nowhere else to go. The explanation of that is that Britain has neglected its technical culture for most of the past four decades. The result is that, against market logic, too many technical people are undervalued (and badly paid) and that there are three million people unemployed.

Some of what the present British government is attempting is meant to redress the balance, but backing an ambitious space programme would plainly work the other way. So the government was right to say no to Gibson. Many other Western European governments, Belgium and Portugal for example, would be compelled, if asked, to say the same. Two or three decades from now, when it may be clear whether Britain's economy will revive, will be time enough to repeat Gibson's question. □

Britain ready to open a telephone line to Japan

- British and Japanese companies competing
- Cable and Wireless breakthrough expected

Tokyo

A dogged effort by Britain's Cable and Wireless to break into Japan's international telecommunications market seems set to succeed. Last week, the Ministry of Post and Telecommunications and Prime Minister Yasuhiro Nakasone indicated willingness to consider separate licence applications from two competing consortia — one headed by the British company, the other made up entirely of Japanese companies. The decision came after efforts to merge the competitors into a single international consortium fell through.

It is more than a year since Cable and Wireless and C. Itoh and Co., a leading Japanese trading company, announced plans to establish an international telecommunications company, International Digital Communications (IDC), to compete with the recently privatized monopoly Kokusai Denshin Densha (KDD) (see *Nature* 321, 803; 1986). At about the same time, another consortium headed by the rival trading companies Mitsui and Co., Mitsubishi Corporation and Sumitomo Corporation formed International Telecom Japan (ITJ), to prepare for entry into the market.

But from the outset there was opposition to the Cable and Wireless group. At first, the Ministry of Post and Telecommunications and KDD insisted there was no room for a third company — an argument belied by KDD's own projections which forecast a tripling in the

value of the market by 1995. Then the Federation of Economic Organizations, with the ministry's blessing, proposed a merger of ITJ and IDC that would have slashed Cable and Wireless's equity holding from 20 per cent to 3 per cent. The ministry argued that it was against "national security" interests for a foreign company to have a controlling interest in a telecommunications company. But protests and threats of retaliation by the US and British governments led to a retraction of the proposed equity holdings.

The ministry, however, continued to insist on a merger of ITJ and IDC, although the two companies have totally incompatible plans: IDC intends to lay its own trans-Pacific optical fibre cable, while ITJ wants to lease circuits on cables to be laid by KDD and American Telephone and Telegraph.

Merger talks have been held on several occasions over the past six months but with no success. The two groups remain resolutely divided over the cable issue. In a last-ditch effort to reach a compromise, IDC at the end of last month proposed establishment of a separate company to lay and operate the trans-Pacific cable; the various shareholders in the merged company would be free to invest (or not invest) in the proposed cable company as they wished. But last week ITJ formally rejected this proposal.

Prime Minister Yasuhiro Nakasone told British Prime Minister Margaret Thatcher during the Venice summit in June that the two consortia would be allowed to apply for separate licences if merger talks failed, a stance that was confirmed by the Ministry of Post and Telecommunications in talks with the US Federal Communications Commission at the end of June. IDC and ITJ are expected to make separate applications soon.

Cable and Wireless estimate that only about one million of Japan's 45 million telephones have access to direct dialling overseas (access requires application to KDD and takes about a month). One factor slowing expansion of the Japanese market is cost. A telephone call to Britain from Japan costs twice that of one made in the opposite direction. Cable and Wireless' director of special projects, Jonathan Solomon, is confident that IDC will be able to undercut present KDD rate by a third, a move that should stimulate additional market growth.

David Swinbanks

US – Japanese supercomputer accord reached

Tokyo

JAPAN and the United States have reached an agreement to simplify procedures for government procurement of supercomputers in Japan.

The accord, formalized on Friday through the exchange of letters between US trade representative Clayton Yeutter and Nabuo Matsunaga, Japanese ambassador to the United States, is intended to help US manufacturers sell multi-million dollar supercomputers to Japanese public institutes. But it is unlikely that the revised procedures will lead to a substantial increase in sales.

Friday's agreement, which took nine months to negotiate, follows the Japanese government's issue last month of new guidelines that came into effect on 1 August. Under the new rules, public institutions (universities and research institutes) must announce well in advance plans to purchase or lease supercomputers. The new procedures require desired specifications be provided to all potential bidders.

The agreement does not end differences between the United States and Japan over supercomputer sales practices. At the signing ceremony, Yeutter said discount practices of Japanese supercomputer manufacturers remain a "major problem". Several universities in Japan have obtained supercomputers from domestic manufacturers at huge discount.

Of the 79 supercomputers delivered or ordered in Japan, only 11 (all Cray computers) are US-made, even though Cray computers have much better software than Japanese brands and dominate the market in the rest of the world.

But cracks are beginning to appear in the public sector market. In July, Aichi Institute of Technology announced its intention to purchase a \$3.74 million Cray computer. And, as part of a \$40,000 million supplementary budget to boost domestic demand, the Japanese government plans an "emergency import" of two or three US supercomputers (one of which is expected to go to the Advanced Telecommunications Research Institute International in the new academic city under development in the Kansai region of western Japan). But with many Japanese universities already having close links with particular domestic manufacturers a rush to buy US supercomputers cannot be expected. Broad interconnected computer systems, not just supercomputers, are becoming the norm. And once a particular system and software have been installed, it does not pay to change. David Swinbanks

Soviet satellite down

Washington

A SOVIET satellite came down unexpectedly in the middle of the Pacific Ocean at 0727 GMT on 10 August, nine days after it was launched into a polar orbit. According to the US Space Command in Colorado Springs, Cosmos 1971 splashed down about 4,800 km north of New Zealand.

The satellite carried scientific equipment for Earth observations and a radio system for accurate determination of its orbit, according to the Soviet news agency TASS. The US Space Command estimates that the satellite weighed approximately 15 tonnes. Tracking data indicated Cosmos 1971 skipped off the atmosphere once before re-entry. Its terminal velocity was estimated at 27,000 km/hr.

TASS said that the satellite carried no dangerous cargo, and there are no reports of debris striking populated areas. J.P.

Coverage for nuclear power plant accidents goes into limbo

Washington

ON 1 August, the law defining the limits of liability for US commercial nuclear power plant operators expired. The expiration of the Price-Anderson Act will have little if any immediate impact on electric power utilities because the old act will continue to cover existing reactors. But there could soon be fallout from Congress' inaction. Price-Anderson also establishes liability limits for Department of Energy (DoE) contractors. Contracts for the operation of three large national laboratories expire at the end of next month, and the contractor has warned that, without liability protection, they will not be extended.

Last year, Congress appeared close to extending Price-Anderson (see *Nature* 322, 676; 1986), but acceptable compromises suddenly evaporated, and the political battle was put off until this year. Opponents of nuclear power saw the act as a way to make the industry tighten up its safety procedures or stop operating nuclear plants. But supporters of nuclear power have argued that it cannot survive without liability protection.

In the event of a power plant accident, the current version of Price-Anderson requires that the initial round of claims from victims be paid from a \$160 million insurance pool. If damage claims total more than that — a virtual certainty for any significant accident — utilities would contribute \$5 million for each reactor it operated to a pool to pay off additional claims. With 110 reactors now operating in the United States, that limits industry liability to \$710 million, still far below estimated claims for any major nuclear power plant accident.

A version of Price-Anderson that the nuclear power industry supports would keep the same \$160 million insurance level, but raise post-accident contributions to \$63 million per nuclear power plant. This would bring the total liability limit to \$7,090 million. The House of Representatives passed such a version late in July. The Senate's version of Price-Anderson is still stuck in committee, but a bill likely to be debated on the Senate floor next month calls for payments of \$60 million per reactor per accident. Other differences between the House and Senate versions are the duration of the extension (10 versus 30 years) and the way claims are handled after the limits of Price-Anderson coverage are reached.

The nuclear power industry has a certain ambivalence about renewing the Price-Anderson Act. All reactors currently operating or under construction are 'grandfathered' under the limits of liability from the expired act. Pushing for a

twelve-fold increase in potential payments — from \$5 million to \$60 million — runs against the grain. With no new orders for nuclear power plants, there is little rush to welcome this increase. But without Price-Anderson, the industry has no future, as no utility would contemplate construction of a nuclear power plant without a limit on liability.

The immediate concern raised by the expiration of the Price-Anderson Act is how DoE will indemnify its contractors. The first problem comes on 30 September, when contracts with the University of California to run Los Alamos, Lawrence Livermore and Lawrence Berkeley Laboratories expire. Two years ago, DoE and the university agreed to renegotiate

the five-year contracts, but the university has made it clear that it could not expose itself to unlimited liability in the event of a nuclear accident. The university receives \$7 million a year to run the laboratories, but total federal contracts to the three laboratories amount to more than \$2,000 million annually. New versions of Price-Anderson would limit liability at the same level set for commercial nuclear plants, but the payments would come from the federal treasury.

One option being considered by DoE to renew the contracts is to invoke the War Powers Act. This allows the Secretary of Energy to indemnify its contractors in the interest of national security, but there are no specific guidelines on how liability is determined under this law. Price-Anderson is clearly the preferred approach, but there is a good chance that the Senate will fail to act before the 30 September deadline. **Joseph Palca**

South Africa ignores IAEA censure over safeguards

Oxford

THE governors of the International Atomic Energy Agency (IAEA) have abandoned their long-running attempt to persuade South Africa to sign the Nuclear Non-Proliferation Treaty (NPT) and have recommended to its General Assembly that South Africa be suspended from the agency.

The scene for the decision was set in December last year when the agency broke off negotiations with South Africa over safeguards at its new commercial enrichment plant. Although South Africa's existing reactors — two at its nuclear power station at Koeberg near Cape Town and a third research reactor, SAFARI 1 — have always been run under IAEA safeguards, the government refused inspection of the enrichment plant at Valindaba, in the Transvaal, which has the capacity to produce weapons-grade uranium. According to the government, inspection could allow too much to be learnt about the "unique" process employed there. The process was hailed as a breakthrough when announced in 1970 but is widely understood to be only a variant of the centrifuge process.

Fears that South Africa may be ready to manufacture nuclear weapons are not new. In 1977 there was an international outcry after the Soviet news agency TASS warned that a nuclear test was imminent. The claim was backed by France, reportedly on the basis of satellite pictures of extensive construction work in the Kalahari desert. Pretoria denied intent to test nuclear explosives. Later, in 1979, a US satellite detected a mysterious double flash, commonly indicative of a nuclear

explosion, off the coast of South Africa. But US civilian experts concluded that the signal was too different from known blast signals to confirm a test had taken place.

South Africa's refusal to sign the NPT has proved costly. In 1974, the South African Electricity Supply Commission (ESCOM) contracted to supply uranium to be enriched in the United States, before being sent on to France to be transformed into fuel rods, and then returned to South Africa. But in 1978 the US Nuclear Non-Proliferation Act was passed which prevented the export of enriched fuel to any country that did not maintain full IAEA safeguards. By 1982, South Africa was in a 'Catch 22' situation: if it did not supply the uranium to its US contractor, it would have to pay huge penalties for breach of contract, and if it did supply the uranium it could not have it back.

ESCOM chose to supply the uranium, but the contract (valued at R116,384,000) was suspended in 1984. ESCOM had to allow the enriched uranium to be sold on the world market while buying enriched uranium from a Swiss company at black market prices. Losses of R56,782,000 could have been avoided if South Africa had been willing to sign the NPT.

While the government has never advocated acquisition of nuclear weapons, recent developments suggest that it is no longer concerned about international respectability in the nuclear field and is prepared to pay the price of going its own way. With IAEA censure, no government would allow contracts to be signed for the planned second nuclear power station in the eastern Cape Province.

Michael Cherry

Resignation of director of UK space centre causes storm

London

A POLITICAL storm has erupted over the British government's refusal to provide any new money for the national space programme and the subsequent resignation of Mr Roy Gibson, director general of the British National Space Centre (BNSC). The future of several of Britain's domestic projects and participation in international ventures is now in doubt.

Last September, at the government's request, BNSC produced a 15-year plan outlining options for British involvement in space. The plan, which has not been published, required an extra £200 million on top of the £112 million that BNSC currently spends. Thatcher's announcement amounted to a rejection of the plan. A compromise sum of £90 million was also refused. But the final straw for Gibson was the government's refusal to provide £7 million to tide over preparatory research for two European projects on which Britain is collaborating. Two years ago, the European Space Agency (ESA) agreed to embark on pre-development studies on the Columbus programme of participation in the international space station, the Ariane V rocket launcher and the Hermes space plane.

ESA had intended to hold a ministerial meeting in June to seek approval to start development of the programmes. That meeting was postponed until November, partly because the British government had not decided how much money it would give BNSC. ESA therefore requested extra funding from member states to continue the preparatory studies on Ariane V and Columbus until the November meeting. The British government refused to provide the extra money so Gibson felt obliged to resign.

Within ESA, Britain's influence will be even less after November's meeting. If it is agreed to press ahead with the three big-spending programmes, substantial extra investment will be required from the member states over the next decade. Britain at present contributes 12.6 per cent of ESA's total budget, behind France (29.6 per cent), West Germany (22.5 per cent) and Italy (14.8 per cent). Without extra money, Britain's stake will fall to around 5 per cent if the proposals are approved, and it will be forced to decline to collaborate in one or two of the three big projects, and thus lose out on lucrative commercial contracts.

Shortly after Gibson announced his resignation, he was visited by ESA's director general, Reimar Lüst, and the chairman of the ESA council, Henrik Grage, ostensibly to discuss ESA's long-term programmes, but they also asked

Gibson to reconsider his decision to resign. Afterwards Lüst made it clear that ESA would push on with its space programme regardless of Britain's participation.

Thatcher's suggestion that the private sector should put up more cash for research has infuriated industrialists, who say that many of the space-exploiting technologies are at such an early stage of development that it will be several years before financial returns permit substantial investment in basic research.

Gibson's resignation has also put in doubt the future of BNSC as an independent body. It was set up two years ago to coordinate Britain's fragmented space policy and to distribute the space funds of the Department of Trade and Industry (DTI), Ministry of Defence and Department of Education and Science. BNSC assumed that it would eventually be given its own money direct from the Treasury. The decision to establish a national space



Roy Gibson, ex-director general of BNSC.

centre was perceived as a government commitment to Britain's future in space; the appointment of Gibson, an internationally respected figure in space research and ESA's first director general, was seen as further evidence that the British government was taking space seriously.

BNSC's management board will meet at the beginning of September to decide the future of the centre. Gibson will leave at the end of that month, to be replaced by his deputy, Jack Leeming, until he retires at the end of the year.

It is possible that BNSC as a body will be disbanded and coordination of the national space programme transferred to the DTI. Sir Geoffrey Pattie, the former minister of information technology who took the lead in promoting BNSC, said last week that Britain had "abdicated from space".

Simon Hadlington

Cold feet in UK over contract financing

London

THERE are indications that the British government is getting cold feet about its controversial plans for financing universities by contracts, after overwhelming opposition to the proposal. A Department of Education and Science (DES) spokesman confirms that the inflammatory word "contracts" may not even appear in the forthcoming Education Bill.

The concept of contracting, which first appeared in the government's white paper, *Meeting the Challenge* (see *Nature* 326, 531 & 532; 1987), and later in a consultative document, briefly spelled out three options: a single comprehensive contract covering all facets of university activities, including teaching and research; separate contracts for each aspect of a university's activities, including individual courses; or a combination of these two, with a core contract covering teaching and various subsidiary contracts.

Greater accountability was a stated goal of contracting, with the underlying intention to direct education and research into areas required by industry. But educationalists argued that the scheme was unworkable, and a threat to the university system. Opposition to a contract system has been virtually unanimous, with every vice-chancellor in Britain speaking out against it in graduation speeches, annual reports and responses to the consultative document. Even some civil servants are said to consider the proposal clumsy and overbureaucratic.

The DES, the University Grants Committee and the Committee of Vice-Chancellors and Principals are due to meet for discussions in autumn, but regardless of the outcome there are signs that a contractual system may not be included in the Education Bill. A DES spokesman says details about university funding are not something for primary legislation, and can be sorted out by the new University Funding Council, which will replace the UGC, "without representing a loss of face or a climbdown".

University sources are pleased, albeit somewhat surprised, that the consultative document on the contract system seems to have been just that.

Kathy Johnston

● The new Education Bill will give business and industry the controlling voice over Britain's 400 further education colleges. The Department of Education has released a report specifying that half the 24 governors on a college governing body will be representatives from employers, and the governors will control college budgets and staff appointments. The government's move continues the trend towards reducing local education authority control over schools and colleges. □

West Berlin to be site of new West German AIDS centre

Munich

WEST Berlin will be the site of the new West German AIDS (acquired immune deficiency syndrome) centre, which will open officially next January. The West German Federal Health Office (BGA) says that the centre will serve as a clearinghouse for information about the spread of the disease as well as a research institute.

The new AIDS centre will be housed in the Robert Koch Institute, itself a division of the BGA, but will operate independently. Its main tasks will include collecting data on AIDS epidemiology, standardizing diagnostic tests for HIV (human immunodeficiency virus), managing AIDS research groups already established in West Germany and conducting basic research into both the 'natural history' of the disease and into its psychosocial effects. Perhaps most importantly, the centre will be given the task of assessing AIDS policy in the various *Länder* (states) and making recommendations to the West German government on how to proceed. The institute is expected to need roughly DM5 million in startup costs in 1988.

West Germany's Research Ministry (BMFT) has spent DM25 million so far on research into the causes and diagnosis of AIDS and is prepared to spend much more next year. The BMFT is placing special emphasis on clinical research into AIDS, apparently to help shore up this traditionally weak area of German research. Pending approval by the Bundestag, the Health Ministry will spend DM135 million on education, counselling and other AIDS-related services, an increase of DM17 million over 1987.

West German health officials have a promising, if unusual, new plan to determine how quickly HIV is spreading to the

population at large. Meinrad Koch of the BGA criticized the suggestion that everyone who is admitted to the hospital should be tested. "Those at risk are seldom in the hospital", he said. Instead, Koch revealed, the BGA is considering carrying out anonymous tests on the blood of trauma patients — those injured in motorcar or other accidents — in order to see whether AIDS is spreading to the heterosexual population. Heterosexuals are visiting West Berlin AIDS clinics more and more often, said Koch, but few of them are HIV-positive.

The West German *Land* of Bavaria is the source of continuing conflict over AIDS policy (see *Nature* 327, 264, 1987). Peter Gauweiler (Christian Social Union,

CSU), state secretary in the Bavarian Interior Ministry, succeeded in introducing a controversial and aggressive AIDS-testing plan there in May. The plan calls for "isolation" of some AIDS carriers and for the testing of all immigrants to Bavaria from non-Western European countries.

The CSU is pressing for acceptance of the plan on the federal level. In July, the CSU delegation to the Bundesrat (the upper house of the federal parliament) introduced a bill that demands that West Germany register all AIDS carriers who act in an "unreasonable" way (such as infected prostitutes who do not use condoms) and that the government be empowered to expel foreigners proved to be HIV-positive. Officials like Koch and West German Health Minister Rita Süssmuth oppose the plan, which, said Koch, is "uncalled for, unnecessary and will not in any way inhibit the spread of the disease".

Steven Dickman

US Congress debates legislation in response to AIDS

Washington

AIDS (acquired immune deficiency syndrome) was very much on the mind of Congress last week as it wrestled with several pieces of legislation outlining strategies for dealing with the disease.

Congress appears willing to commit large sums of money for programmes to test for HIV (human immunodeficiency virus), the virus causing AIDS, but the debate has focused on whether such tests should be voluntary.

There are some eight different bills before the House of Representatives that address the testing issue. The one that appears to have the broadest support has been introduced by Henry Waxman (Democrat, California), chairman of the health subcommittee of the Energy and Commerce Committee. Identical legislation has been introduced in the Senate.

Waxman's bill would give the Department of Health and Human Services \$400 million a year for the next three years to establish centres offering voluntary screening for HIV antibodies. The bill requires that testing centres must offer counselling both before and after the test is given and after the results are known. The legislation also lays down strict confidentiality provisions, with fines for violations of privacy. The new act would also prohibit discrimination against those who have evidence of antibodies to HIV.

A stream of public health professionals representing the pillars of the health community — the American Medical Association and the American Nursing Association among others — paraded before Waxman's sub-committee to endorse his legislation.

But sitting to Waxman's left in the committee room was William Dannemeyer (Republican, California), who has different views about how testing should proceed. Dannemeyer has introduced seven separate pieces of legislation that would require mandatory HIV testing for certain segments of the population — couples seeking marriage licences, prostitutes and hospital patients. Dannemeyer also favours disclosing HIV-antibody status of individuals who might infect others, and paraded his own stream of witnesses.

Although voluntary testing and confidentiality are supported by most public health officials, including those in the US Public Health Service, some in the White House lean toward Dannemeyer's view that the civil liberties of a minority of the population may have to be compromised to protect the health of the majority.

The House of Representatives also passed last week a bill authorizing the creation of a National Commission on AIDS. The 15-member commission, appointed by the House, the Senate and the president, would consist of at least eight members from the scientific and medical communities qualified to serve "by reason of education, training or experience". The commission would be charged with advising Congress and the president on policy priorities in research and public health. The House moved with unaccustomed speed to pass the bill creating the commission, in part as a response to criticism of the president's commission announced two weeks ago (see *Nature* 328, 372; 1987). It is not yet clear when the measure will be brought up in the Senate.

Joseph Palca

Symbolic computer institute in Linz

Munich

THE University of Linz in Austria has announced the founding of a new Research Institute for Symbolic Computation (RISC-Linz) with the help of Austria's nationalized steel company VOEST-ALPINE. The institute, which is also supported by the Austrian and local governments, will investigate all areas of computation that use non-numerical symbols, such as logical formulae and geometric objects. The research at RISC-Linz, which is related to artificial intelligence but has a more industrial bent, will be carried out by a core staff of seven faculty members from Linz University as well as five researchers from industry. S.D.

Glasnost lost in French government's science policy

Paris

THE Conseil Supérieur de la Recherche et de la Technologie (CSRT), established by the previous French socialist government with a legal right to be consulted on civil science and technology budget proposals, is being bypassed by the present Conservative government. For the second year in succession, the CSRT will not be shown the budget until after the proposals have been put to parliament next month, defeating the purpose of consultation, emasculating the committee and possibly infringing a 1982 decree.

The CSRT was set up in 1983 under the chairmanship of the then research and industry minister, Jean-Pierre Chevènement, ostensibly to enable his ministry to shape its policy to suit the long-term needs of science and industry. A practice of open debate was inscribed in the CSRT's legal constitution, with the 40-strong membership of academics and industrialists encouraged to play a major and public role in the summer run-up to the annual budget debate. Although it endorsed plans for radical reform in French science and technology, the CSRT

held the view that fundamental research would need to grow steadily to meet future unpredictable technology needs.

With the change in government in March 1986, and a shift in emphasis from public sector to industrial research, the CSRT found itself increasingly excluded from the corridors of power. Neither the 1987 science budget nor the controversial university reforms, which ended in the resignation of their initiator, education and research minister Alain Devaquet, were discussed with the CSRT. Touchstones of the previous relationship between government and CSRT, such as the legal commitment to see recruitment of young scientists to the state research centres grow by 1,400 per year, have also been abandoned. Some forecasts predict that no new research posts will be created this year in the principal research centres of the Centre National de la Recherche Scientifique (CNRS).

The diminishing voice of fundamental research is also reflected in the composition of the new CSRT committee, with a two-year term of office, named last month by Minister for research and higher education, Jacques Valade. Jean-Pierre Causse, who comes from industry, replaces Françoise Kourilsky as executive vice-chairman of the committee. Valade's influence on the research budget has also been dramatically reduced. Last year, 75 per cent of the research budget was transferred to the industry ministry. As the CSRT was set up to advise the research minister, its power has effectively been reduced in direct proportion.

According to a CSRT report*, which has not been endorsed for publication by the government, a change in the way the research budget is calculated has enabled the government to sidestep its legal obligation to involve the CSRT in the impending budget debate. The 1982 decree requires consultation on the budget for civil research and technological development (BCRD), which was calculated by pooling research expenditure forecasts from different ministries, but excluded defence and salaries for university dons. In 1986, the incoming government created a new budget reference, the 'budgetary effort for research and development' (EBRD), which includes spending on defence and post and telecommunications. Because the BCRD is now an *a posteriori* abstraction, worked out from the EBRD, the legal obligation to consult the CSRT has become meaningless.

In the past, 'indiscretions' on the part of the CSRT gave the press a foretaste of the autumn science budget as early as July. This year it seems that the CSRT, as well as the public, will have to wait until September.

Peter Coles

New York launches test programme for AIDS virus

Washington

NEW YORK, the US state with the second highest incidence of acquired immune deficiency syndrome (AIDS), will be the first to take steps on its own to assess the spread of the AIDS virus within its borders. Last week, Governor Mario M. Cuomo announced that the New York state health department will screen 100,000 blood samples selected randomly from routine specimens taken at state hospitals for antibodies to human immunodeficiency virus (HIV), the virus causing AIDS.

The \$3.4 million study will give public health officials an indication of how far the virus has spread through different sub-populations and regions of the state. This will allow them to assess the burden the disease will place upon health care and insurance systems, and help to target preventive educational campaigns. Cuomo cautioned against "grasping for simplistic answers to complex questions" or "believing . . . that if everyone were simply tested for the virus, that that in itself would stop it", and stressed the importance of a scientific approach to studying the spread of the disease.

The tests will be anonymous, and patients will not be notified of the results. The samples used in the study will be accompanied only by clinical information that will enable researchers to determine

if the donor fell into a high-risk group, such as homosexuals or intravenous drug abusers. Testing will begin in October, and is expected to take six months to a year to complete.

The New York study will be one of the first nonmilitary sampling programmes in the US. The Department of Defense has been screening its personnel since 1986. But HIV frequency in the military is unlikely to be representative of the population at large.

The Centers for Disease Control (CDC) set up a regional programme very similar to the New York study in late 1986. Since then, four "sentinel hospitals" have been monitoring 300 anonymous specimens a month for antibodies to HIV.

According to Timothy Dondero, the director of the study, testing anonymous samples to construct trends and levels of HIV prevalence is the "only ethical way". First announcement of results will come after Dondero has six months of data.

The protocol for the nationwide random survey for HIV, announced by Secretary of Health and Human Services Otis Bowen at the Third International Conference on AIDS in June, is still being developed. The national survey is intended to test the blood of 45,000 people in order to estimate the number who are infected with the virus in the United States.

Carol Ezzell



*4 *Années de Conseils*. Conseil supérieur de la recherche et de la technologie. March 1987.

West German release of altered bacteria causes furore

Munich

THE release of genetically manipulated bacteria onto a pea field in West Germany has set off a sharp debate and put scientists and public officials on the defensive.

The bacteria are of the genus *Rhizobium*, which adhere to the roots of leguminous plants and fix nitrogen. They were manipulated *in vivo* to include a transposon containing an antibiotic resistance gene intended to serve as a genetic marker. The releases, made somewhere in Bavaria in May, first came to public attention in West Germany in late July.

The experiment was carried out by Walter Klingmüller, a geneticist at the University of Bayreuth, under the auspices of the Biotechnology Action Programme of the European Economic Commission (EEC). It followed similar studies conducted at the Rothamsted Experimental Station in Britain and at the National Institute for Agricultural Research in Dijon, France.

The purpose of the experiments, said Klingmüller, was threefold: to determine how many of the bacteria survive; to determine whether any of the genetic material, that is, the transposon-containing plasmid, is transferred from the altered *Rhizobium* to other bacteria; and, in the event that the plasmid is transferred, to determine whether the antibiotic resistance gene remains in the bacteria's DNA even if the plasmid itself breaks down.

According to Klingmüller, the first set of results indicate that most of the altered *Rhizobium* have, not surprisingly, died off. It has not yet been determined whether the transposon was transferred to other bacteria. But the furore that followed public disclosure of the experiment had more to do with the official attitude towards the release than it did with the results. In order to comply with West German regulations on genetic engineering, any experiment involving the release of "genetically manipulated organisms" must be approved by the Central Committee for Biological Safety (ZKBS), a division of the Federal Health Office.

But "genetically manipulated organisms" are defined to include only those created *in vitro* using recombinant DNA techniques. The *Rhizobium* released in Bavaria do not fall under this rubric and therefore needed no official approval.

This distinction appears nitpicking to the critics of genetic engineering, among them the Green Party and Wolf-Michael Catenhusen (Social Democrat), chairman of the Research and Technology Committee in the Bundestag (parliament).

Catenhusen chaired an Enquete Com-

mission on Genetic Engineering during the last Bundestag session that recommended a five-year moratorium on any release of genetically altered organisms into the environment. That none of the authorities responsible — the ZKBS or the Research Ministry (BMFT), for example — made the experiments public is a "political snub" to the public and the Bundestag according to Catenhusen.

The Greens were also outraged that the ZKBS used the *in vivo/in vitro* distinction to justify the experiments, and find it incredible that the BMFT report of September 1986 stated that there have been no projects proposed to the ZKBS "in which

a proposed goal is to develop genetically manipulated organisms for later release into the environment".

Even before the West German release was announced, a Green representative to the European Parliament, Benedikt Härlin, had introduced a resolution to the EEC to place a moratorium on all releases of altered organisms into the environment. A previous draft proposal had called for regulations only on organisms manipulated using *in vitro* recombination.

Catenhusen is determined to introduce the subject early in the next Bundestag legislative session this autumn. The Deutsche Forschungsgemeinschaft, which funds much of West German basic research, said in May that it rejects a moratorium, although it recognizes the danger of releasing organisms that have been manipulated *in vivo*. **Steven Dickman**

UK publishes report on CFCs based on old data

London

BRITAIN'S Department of the Environment (DOE) has come in for sharp criticism for its fanfare in issuing a report on chlorofluorocarbons (CFCs) and stratospheric ozone based on two-year-old data.

In the first sentence of its press release heralding the newly published report of the Stratospheric Ozone Review Group, the DOE claims that the present rate of CFC production is unlikely to lead to a reduction of the stratospheric ozone layer. But the group's report is based mainly on the findings of the assessment carried out by the US National Aeronautics and Space Administration and World Meteorological Organization in 1985. Ozone research is developing rapidly, and many scientists involved consider that assessment to be out of date.

Some scientists argue that the government should have quietly buried the report and arranged for an immediate update based on the latest findings of the role of chlorine in the formation of the Antarctic ozone hole. "It is very naughty of the DOE to dress up that report with a press release", says Dr Joe Farman of the British Antarctic Survey, a member of the Stratospheric Ozone Review Group.

The report concludes that one-dimensional models, which are subject to large uncertainties, predict that globally averaged column ozone will not be depleted at present levels of CFC emissions and other source gases. It points out that two-dimensional models predict larger decreases in ozone. The report summarizes both types of models as well as experimental evidence.

The DOE's press release also says that advice from the report allowed the government to formulate a policy on

regulating CFCs based on "a sound understanding of the science". But pressure groups maintain that the government wants to use outdated research to bolster the position that there is no need for stringent regulatory controls on CFCs. Britain has been accused of dragging its feet in the international negotiations on a protocol to control emissions of CFCs to protect the ozone layer, although recently its negotiators have shown more willingness to compromise.

There are indications that countries involved in the Vienna Convention for the Protection of the Ozone Layer will finally reach agreement and sign a protocol in Montreal next month calling for an eventual 50 per cent reduction in CFC production. Farman says there is widespread agreement among scientists that an immediate 85 per cent cutback is needed just to hold present atmospheric chlorine levels constant, because chlorine is being pumped into the atmosphere six times faster than it can be removed naturally.

Antarctica's seasonal ozone hole cannot be satisfactorily explained by anything except CFCs, according to Farman (see *Nature* 328, 463; 1987).

The DOE points out that emissions of methane can compensate for CFC-induced ozone destruction. However, scientists argue that chemical engineering of the atmosphere is not a satisfactory solution, especially when greenhouse gases are involved.

Ozone experts are worried that if there is a critical level of chlorine in the atmosphere, it has already been surpassed in the Antarctic; they argue that if CFCs continue to be emitted, the critical levels could be surpassed in middle latitudes too.

Kathy Johnston

Unilever purchases British plant breeding interests

London

In a move aimed at extending its growing agricultural interests, Unilever, the Anglo-Dutch consumer conglomerate, is to pay the British government £66 million for the plant breeding assets of the Plant Breeding Institute (PBI) and the National Seed Development Organization (NSDO), which markets new varieties produced by PBI. The deal will be completed by 30 September this year.

The PBI/NSDO package attracted widespread interest from the food, chemical and agricultural industries; after

the sale was announced in May (see *Nature* 327, 176; 1987), the government received "dozens" of applications for the confidential memorandum of information, reserved for "serious" bidders. Unilever was shortlisted with Imperial Chemical Industries (ICI) and the food group Booker. Of the three, Unilever's bid was the highest by a considerable, though unspecified, margin.

The plant science activities of PBI were not included in the package, but will be retained by the Agricultural and Food Research Council (AFRC) and relocated in 1990 to the new Institute of Plant Science Research. The decision to hive off the scientists provoked hostility from within PBI, which complained that the close interaction between the molecular biologist and the breeder was fundamental to the institute's success. More than 80 per cent of NSDO's revenue is derived from royalties on varieties produced by PBI; in 1986, PBI/NSDO generated income of £11.3 million with an operating profit of £4.5 million. In the same period, PBI received £3 million in government grants, with a further £1 million in contracts. The proceeds from the sale will go to the Treasury.

Given that the sale had been decided, staff at PBI felt that the terms being offered by Unilever were more favourable than those of the other two shortlisted bidders. Unilever intends to allow PBI (to be renamed PBI Plant Breeding) to retain a degree of autonomy and will move its own pea- and bean-breeding programmes to PBI's Cambridge site. PBI's breeding programmes, for wheat, barley, potatoes and field beans, will be retained.

Most of Unilever's plant science research is carried out at Colworth in Bedfordshire, where it is best known for its work on cloning oil palms. Its current agricultural activities focus largely on plantation crops. The Colworth laboratory is also carrying out research on the genetic manipulation of oilseed rape. Although PBI has a rapidly expanding oilseed rape breeding programme, this was not included in the sale. Instead it will go to the Agricultural Genetics Company (AGC) in settlement of various claims that AGC has on PBI. AGC, set up in 1983, was given first option to exploit biotechnological discoveries made by AFRC.

Last year, Unilever's agricultural activities accounted for £632 million of the group's total turnover of £17,140 million. It produced an operating profit of £35 million, of a total £1,124 million. Unilever's overall research and development budget was £305 million last year, compared with £253m for 1985.

Simon Hadlington

SDI plans play safe with Congress

Washington

THE US Strategic Defense Initiative Organization (SDIO) appears to have gone out of its way to avoid confrontation with congressional critics in its new plans for a series of tests of anti-missile interceptor rockets and tracking satellites. All tests have been designed to comply with a strict interpretation of the Anti-Ballistic Missile (ABM) Treaty.

The new plans were revealed last week in a series of environmental impact assessments that the Department of Defense is required to make before undertaking major operations. Thirteen large-scale tests are planned for the next five years and paint a picture of a missile defence system that is based largely on conventional technology. Plans to test exotic laser and particle-beam weapons are left for the future, while emphasis is placed on building on successes already scored in intercepting incoming missiles with high-acceleration rockets.

For the first time, all the major components of the system are to be tested, including satellite surveillance and command, ground-based surveillance and tracking, battle management, control and communications, and two types of 'kinetic-kill' interceptor rocket. One, the exoatmospheric reentry vehicle (ERIS) will be carried aloft on a large missile and will try to hit incoming warheads before they can re-enter the atmosphere. The other, the space-based interceptor (SBI) is intended to destroy missiles much earlier. In its final configuration, SBI homing rockets would be fired from an orbiting space platform and collide with missiles as they leave the atmosphere at the end of their boost phase. But in the planned tests, the ABM Treaty ban on testing "space-based" systems will be observed and the homing rockets fired from the sub-orbital flights of a large booster rocket.

The tests will be extremely expensive to perform. From four to seven Minuteman or Polaris rockets will be launched from California to provide targets for ERIS missiles fired from the Kwajalein Atoll in the Pacific Ocean. Smaller sounding rockets will be used to test the SBI homing rockets intercept capability at the Kwajalein Atoll. A new satellite will be also be needed to test space-based surveillance and tracking systems. Opposition to the plan may still come in the SDI review under way at the Department of Defense and from Congress. Earlier attempts in the Senate to restrict SDI testing (see *Nature* 327, 175; 1987) ground to a halt as the Irangate hearings got under way but are expected to be revived again in September.

Alun Anderson

AIDS programme head

London

DR Geoffrey Schild, director of the National Institute for Biological Standards and Control (NIBSC), has been appointed director of the Medical Research Council's (MRC) directed programme of AIDS (acquired immune deficiency syndrome) research from October.

Schild, 51, will divide his time between directorial duties at NIBSC, which has recently been designated a World Health Organisation collaborating centre on AIDS, and the MRC programme. He succeeds Sir James Gowans as head of the directed research programme, begun early this year.

MRC has £14.5 million to allocate for research into vaccines and antiviral therapy, with some projects already under way.

K.J.



Soviet superconductors

London

THE Soviet Union is to produce a standard high-temperature superconducting ceramic in sufficient quantities to supply all Soviet laboratories working on superconductivity. Following the production of the first Soviet high-temperature superconductor earlier this year in Sverdlovsk, all interested laboratories began producing their own materials. As the superconducting properties of different ceramics differ considerably, the production of a standard material will allow experiments carried out in different parts of the country to be compared and correlated. Bulk production of the material is to begin at once.

V.R.

Boycott of South Africa

SIR—As academics born and bred in South Africa, we are distressed that eight pages of *Nature* (327, 269–276; 1987) have been dedicated to opposing the boycott of apartheid science and technology.

There are two fundamental flaws in Maddox's arguments for continued support for South African science and technology. The first is that he appears to see the privileged white minority, rather than the oppressed black majority, as the principal agents of change. The second is that he fails to make the crucial link between scientific and technological education and research and the maintenance of apartheid.

Symptomatic of the first error is that he seems to have met only representatives of the white scientific and academic community and directors of corporations with a vested interest in the system, with the single exception of the University of the Western Cape, which has 10,000 students, mostly coloured. Significantly he reports that this university "happens to be in favour of an academic boycott of South Africa but selectively". It is, perhaps, not surprising that he did not meet more black scientists, given that the very essence of apartheid is to deny opportunity for educational advancement to the black majority.

Maddox acknowledges that the boycotts are "biting hard" on whites, and that they are making the "technological community in South Africa reflect more carefully about its domestic political environment". What he misses is that the principal objective of sanctions is to give the maximum possible reinforcement to domestic anti-apartheid forces. They are, in fact, the principal agents of change.

Curiously, he acknowledges the tenacity with which the white population clings to apartheid (as demonstrated again in the recent election) yet he believes that after decades of failure to bring down apartheid by scientific contracts, science can now be some kind of "Trojan horse".

Maddox also claims that it is the view of "liberal South African critics of apartheid" that "one person, one vote, tomorrow would indeed be a recipe for disaster". That is certainly not the view of the overwhelming majority of black South Africans, nor of progressive whites, for whom universal adult suffrage in a unitary and democratic state is their fundamental goal.

The second point is even more serious. Foreign science and technology have been crucial in South African economic development, in fact even more important than foreign capital. This is implicitly acknowledged so far as military technology is concerned, in the statement that "conferences with a military significance

might be closed to South Africans". South Africa is nowhere near to scientific and technical self-reliance. Maddox appears to agree with a ban on the sale of microchips and other technology, but illogically draws the line at sanctions against the transfer of technical information and the training of skilled personnel. It is clear that he wants to give every assistance to "help to bridge the emerging shortage of skilled people".

The answer to this is given by Richard Moorsom, an independent researcher, in his recent book *The Scope for Sanctions* (Catholic Institute for International Relations, 1986), who points out that

"the racism of apartheid makes it peculiarly vulnerable in this crucial area: the talent pool is restricted to the white minority, who must also staff the large state apparatus of civil and military repression as well as the big parastatal corporations. Consequently, skilled expatriate staff are an essential link in the transfer of technology by the transnational corporations. Regulations applied to companies could discourage them from servicing or collaborating with South African companies, parastatals or state agencies. Included would be any support for the build-up of local research and developmental capacity...."

"Easier to restrict would be South African access to research and training programmes in the industrial countries themselves...."

It is also misleading to say that the scientists affected by sanctions "merely live" in South Africa as if it were some kind of planet from which they could not escape, or in which they could not actively join the struggle against apartheid. We know several South Africans who could have had distinguished scientific careers but have chosen instead to face imprisonment and hardship. Many others have emigrated. Those who remain and, at best, limit themselves to tepid condemnations of apartheid, pay lip-service to democracy but in fact contribute to the survival of apartheid.

Maddox argues that "the exclusion of South African scientists [from a conference] ... was absurd because those concerned were bound to have particularly interesting things to say". Pursuing his analogy with the boycott of South African rugby, which he admits has had noticeable effects, one could argue that one should not boycott the Springboks because they play so well. It is because they have something to say, and something to learn, that sanctions hurt those scientists and technologists whose work sustains apartheid. The argument for sanctions is utilitarian: the relatively small costs to the world scientific community of forfeiting white South African participation are far outweighed by the enormous benefits that will accrue to South Africa and the rest of

the world by hastening the process of change.

The ending of links with South African science and technology is a vital part of the wider movement for the total economic, military and cultural isolation of the apartheid system. The Commonwealth Eminent Persons Group said that concerted action of this kind "may offer the last opportunity to avert what could be the worst bloodbath since the Second World War". Academics have a special responsibility to distance the tenets of international science from the self-interest of those who enjoy the privileges of white minority "science".

HILLIARD FESTENSTEIN
JOHN SACHS

*Department of Immunology,
London Hospital Medical College,
Turner Street,
London E1 2AD, UK*

BOB HEPPLE

*Department of Laws,
University College London,
Gower Street, London WC1E 6BT, UK*

SYDNEY SHALL

*Cell & Molecular Biology Laboratory,
Biology Building,
University of Sussex,
Brighton BN1 9QG, UK*

Better refereeing

SIR—Gideon Gilat recently proposed an interesting scheme for improving scientific refereeing (*Physics Today* 40, 147; 1987), based on the publication of controversial papers with an optional right for a referee to have his or her signed comments published alongside. I have an even simpler suggestion which will not result in increased correspondence for the editors. When the manuscript is sent for review, the cover letter should say that "the referee is asked to review the paper on the understanding that the editors reserve the right to publish the submitted paper in its original form and the full signed texts of some or all reports of the referees".

This option, which in practice may not be exercised too often, will undoubtedly not discourage the main body of referees from giving an honest and objective critique. At the same time, it will create an effective defence against reports driven by, for example, jealousy or fear of competition. As a first step, the editorial boards of major journals could consider the adoption of such a policy (or that proposed by Gilat) on a trial (term) basis.

ALEXANDER A. BEREZIN
*Department of Engineering Physics,
McMaster University, Hamilton,
Ontario, Canada, L8S 4M1*

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only. □

Modelling for its own sake

A new calculation of a neural network model, while mathematically interesting, seems too close for comfort to the domain of modelling for the fun of it.

SIMPLE models of complicated systems have several functions, of which the chief is to give ordinary mortals a comprehensible picture of a phenomenon. Thus the old schoolboy calculations of the bending of a beam under a load are a representation of how a wooden plank may bridge a stream without leading wise people to expect the calculations to apply to real wooden boards, with their fibrous structure. Sometimes, simple models point to general properties that apply even to the complicated systems they represent; Bohr's atomic model gives a good account of how energy levels crowd together near the ionization limit even though its account of the splitting of energy levels in magnetic fields is incorrect. But model-building can get out of hand, as with the fruitless refinement last century of the luminiferous aether to accommodate electromagnetic radiation.

Is there a danger that the same may now be happening with neural network models of the brain? This sour question is provoked by an intricate calculation of a particular neural network by Ronny Meir and Eytan Domany of the Weizmann Institute (*Phys. Rev. Lett.* **59**, 359; 1987). The answer is that there is indeed a danger that neural networks will become a field of study in their own right, unrelated to the phenomena they model, but wringing useful notions from them will increasingly require discrimination.

Neural network models are mostly elaborations of a crisp formulation by J.J. Hopfield (*Proc. natn. Acad. Sci. U.S.A.* **79**, 2554; 1982). An earlier article on this topic (*Nature* **325**, 11; 1987) should have made Hopfield's contribution clear. The idea is to represent neurons by two-valued entities — neurons may be either ON or OFF — and the state of each may be influenced by the states of all others. If the neurons modelled are those of the sensory cortex, the state of the system is, at any time, the brain's impression of the outside world. Hopfield's achievement was to show how, by a suitable choice of the parameters describing the mutual influence of the neurons, to embed "memories" in this system. The embedded memory states are orthogonal to each other in the sense that arbitrary impressions of the external world which overlap with, or evoke, predominantly one of the embedded states will not also evoke another. The embed-

ded memory states are a little like the eigenstates of a quantum system, and thus a basis for the analysis of arbitrary states.

This model has been extraordinarily stimulating. It offers an explanation of the distributed character of memory, as when people whose brains are damaged by accident or surgery find a general impairment, not a selective loss, of memory. Similarly, it is a model for the process of learning, represented as that of fixing the mutual interactions of neurons to define particular ground states, and where feedback of various kinds may be especially important (*Nature* **328**, 107; 1987).

All kinds of intriguing speculations also suggest themselves. Is lateral thinking a measure of the overlap between supposedly orthogonal ground states? Or a dynamic phenomenon in which the processing of sensory information evokes a sequence of partially relevant images? If the sensory cortex is continuously responsible for processing sensory information, what structures in the brain monitor those events and tell which memories are instantaneously evoked? Can one remember moving pictures (apparently yes, see *Nature* **325**, 11; 1987)? Will the model work for language-learning and, if so, does the common structure of human language for which Chomsky has argued say something about the permissible eigenstates of that part of the brain? All these are coffee-table questions now made a little more tangible. At this rate, the question "What is consciousness?" will again become respectable.

So it is natural that much energy has been spent on elaborations of the neural network model. Meir and Domany have worked with their own layered neural network in which they simulate the columnar structure of the cortex with a sequence of layers of model neurons, each of which is influenced by all cells of the preceding layer. The model provides a means of transforming an image of the external world at the lowest layer into other representations at succeeding layers. As there is no feedback from upper to lower layers, it is also a model of how a single neural sheet might transform an image of the outside world with the passage of time, in which case the parameters specifying the strength of the interconnections become the dynamic rules for the evolution of the system. On the first layer, the

representation of the eigenstates of the system, helpfully called "key patterns", is externally dictated, but their representation in other layers is entirely independent. The object of the exercise is to calculate the likelihood that an external stimulus more or less congruent with a key pattern on the first layer is itself transformed so as to be congruent with the transformed image of that same key pattern on succeeding layers.

With certain assumptions, in particular that the number of neurons in each layer tends to infinity, the conclusion is that this is a high-fidelity system provided that the number of stored key patterns is not too great a fraction (0.269) of the number of cells in each of the layers. But if this critical fraction is exceeded, correlation vanishes between the successively transformed versions of the input stimulus and the stored versions of the key patterns.

Does this help those who seek a clearer picture of how the brain functions? That a system of layered neurons can be calculated exactly is something to be pleased about; part of the price paid for this exactness is that the representations of key patterns in successive layers are supposed independent of each other, which is probably not the case in the real columnar cortex. Even so, the same mathematical tricks will no doubt be useful in the calculation of later network models. That the number of key patterns that can be stored in such a system must be limited if memory is to function well is interesting, although that was evident in Hopfield's work five years ago.

The sharp phase-like transition, as the ratio of stored memories to active neurons in each layer increases above the critical fraction, from a condition in which memories are faithful to one in which memory completely fails, may suggest to some the more or less sudden onset of memory impairment in conditions such as Alzheimer's disease; declining numbers of active neurons allow the ratio to exceed some critical value. On that view, the underlying trouble is just as much the excess baggage of stored key patterns as the decline of the number of active neurons. But that, of course, is unbridled speculation, of which Meir and Domany's paper is innocent. That is why it seems uncomfortably near the borderline of modelling for its own sake. **John Maddox**

Cerebral cortex

A quiet revolution in thinking

Jennifer Altman

WE ARE still far from understanding how the cerebral cortex functions, but a promising shift in ways of thinking about cortical organization emerged from the discussion at the 50th Dahlem conference*. The long-held view that the cortex is a hard-wired neuronal machine with its operations shaped mainly by evolution and development is giving way to a concept of basic processes continually modified by feedback and lateral interactions. Such modulation could provide the plasticity necessary for adaptive behaviour and learning.

Why has it taken so long to recognize this plasticity in cortical operations? Perhaps the regularity and seeming homogeneity of cortical anatomy has been misleading, creating the impression of an almost crystalline matrix that is set up during development and then is immutable. Perhaps, too, the complexity of the cortex has forced people to focus on local questions without considering the wider implications. The strongly reductionist approach has not helped as most experiments are done in biologically impoverished situations, either in anaesthetized animals, using precisely defined but non-biological stimuli often in inappropriate contexts, or with animals that are highly trained to perform stereotyped tasks. The convergence of evidence from anatomy, physiology and behaviour is now highlighting these shortcomings.

Functional maps

The cortex essentially makes a map of the world, but what is mapped and how the maps are used are still much debated. Over the past 20 years, the use of modern anatomical tracing methods and physiological recording has established that each sensory modality is mapped several times in different areas, with about a dozen representations of the visual world and half a dozen each of auditory inputs and body sensation (John Allman, California Institute of Technology; Jon Kaas, Vanderbilt University). The maps differ in the attributes of the stimulus represented, in how the field is emphasized and in the types of computations performed.

Clearly, the specificity of all these representations means that cortical maps can no longer be considered simply as projections of the receptor sheet; instead it is functions that are mapped (Allman; Semir Zeki, University College London). This idea is reinforced by the recent demonstrations that single areas can

contain multiple, interdigitating maps, each representing a different attribute of the stimulus. In the first two visual areas, for example, groups of cells selective for wavelength are found among columns of cells selective for orientation and/or direction of stimulus movement¹⁻³.

This separation of different attributes of the stimulus into parallel processing channels has now been shown throughout the monkey visual pathway (Zeki) and in the auditory cortex of the moustache bat (Nobuo Suga, Washington University, St Louis). Each channel, or information stream, can be seen as dealing with a particular question about the world. In the visual pathway, separate coding of size, shape, colour, position and direction of movement of an object begins in the retina. In the thalamus and lower levels of the visual cortex, the responses of neurons to these attributes are increasingly selective and more derivative. Direction and spatial location — where is the object? — are then channelled mainly to the parietal lobes, form and colour — what is it? — to the temporal lobes (ref. 4; see figure). These in turn project separately to the prefrontal lobes, which are involved in the short-term maintenance of information required for planning actions. Here, Patricia Goldman-Rakic (Yale University) has identified one area concerned with delayed-response tasks, which require remembering where an object is, and a second for object-matching tasks, requiring knowledge of what objects are. On the output side, the questions are slightly different and Giacomo Rizzolatti (University of Parma) identifies separate maps in premotor cortex, one dealing with the planning and initiation of actions, the other with their guidance and execution — what to do and how to do it.

The number and organization of the maps varies between species, reflecting what is of behavioural importance to each. At lower levels of the cortex, the maps bear a fairly simple topographical relationship to the world but in higher areas, precise topography may be sacrificed for the mapping of more abstract functions. Here, selected aspects of the input are combined in ways that are likely to be relevant to a behaving animal. In the bat auditory cortex, for example, Suga showed how maps that plot the velocity and range of the prey, its position relative to the bat, and how it moves are computed in three-dimensional spatial coordinates. In the posterior parietal cortex in monkeys, information about eye position relative to the head and to the position of

the image on the retina are integrated for locating objects with respect to the position of the head in space (Richard Andersen, Massachusetts Institute of Technology). Neurons in this area have very large overlapping receptive fields that are modulated by the angle of gaze, so that the strength of the responses of a neuron to an image at a particular location on the retina varies according to the position of the eyes. An underlying topography may, however, be maintained by reference back to the more topographically organized areas that provide the inputs.

Networks

Andersen caused considerable excitement at the conference with his description of a neural-network model of this area he has developed with David Zipser (University of California, San Diego). This uses a back-propagation algorithm⁵ to train a three-layered network to learn spatial positions. In less than 1,000 trials, the outputs of the network give accurate positions in head-centred coordinates for particular combinations of retinal-image position and two eye positions. Remarkably, the intermediate elements in the network develop receptive-field properties that seem very similar to those of the neurons in the posterior parietal area. The model confirms that a topographical representation is not required to provide readout of spatial location and, more important, indicates that the neurons in this area probably learn the association between eye and retinal positions. It also shows that an accurate spatial representation can be set up with very few eye positions, as seems to be the case in the posterior parietal area.

The Andersen–Zipser model is one of the first applications of neural-network theory to a neurobiological problem. Although some at the meeting were sceptical that the approach could be generally useful, this example has also predicted other response properties of the neurons that have subsequently been found. More generally, it reinforces the idea that neuronal operations can be probabilistic.

It is not known how the information streams are recombined to provide the perception of objects as entities or how the readout works. It is now thought unlikely that the streams converge on to single object-specific 'cardinal' or 'grandmother' neurons. More probably the perception of an object results from continual cross-reference between the streams so that activity in one is shaped or gated by that in others. At the same time, the integrity of information in the individual streams must be preserved, because we know that the component attributes of an object, its colour, size, shape and so on, can be reported separately. The extensive connections between the temporal and

**Neurobiology of Neocortex*, Berlin, 17–22 May 1987.

parietal lobes are probably part of the cross-reference system, though curiously, anatomical tracing experiments show that where projections from two areas converge in a third, they never overlap but always terminate in adjacent patches (Goldman-Rakic).

It seems likely that a single perception results from a specific combination or ensemble of neurons located in several areas, each responding to a particular feature of the stimulus field. Even the neurons in the temporal lobes of monkeys that respond to complex stimuli such as views of faces or of particular body movements⁶⁻⁸ do not each give a complete account of the stimulus and it must need an ensemble of them to achieve specific recognition (Charles Gross, Princeton University). Memories could be consolidated by mechanisms that facilitate the coordinated activity of such an ensemble once one is established.

Even the idea that each attribute of a stimulus is coded in a series of hierarchical steps of increasing abstraction is being questioned. There is now evidence that return connections from higher to lower cortical levels are important in shaping responses and filling in the fine detail not discretely represented at higher levels. New work from Adam Sillitoe's group (University College, Cardiff) provides a nice illustration. Neurons in primary visual cortex and its thalamic input area, the dorsal lateral geniculate nucleus, respond to bars of defined lengths. Sillitoe's group has shown that this tuning results from interactions between neurons in the two areas and so cannot be separately attributed to either.

Long-range lateral connections within an area may also be important for shaping the response selectivity of neurons and for modulating responses according to the context of the stimuli. Charles Gilbert (Rockefeller University) reported precise connections in the primary visual area between groups of cells with similar colour responses and there may also be specific connections between orientation-response cells. The role of reciprocal connections, both within and between areas, will be an important focus of attention in the near future and it seems likely that neurons at all levels in the processing pathways will turn out to be part of the ensemble that recognizes, and perhaps remembers, a particular stimulus constellation.

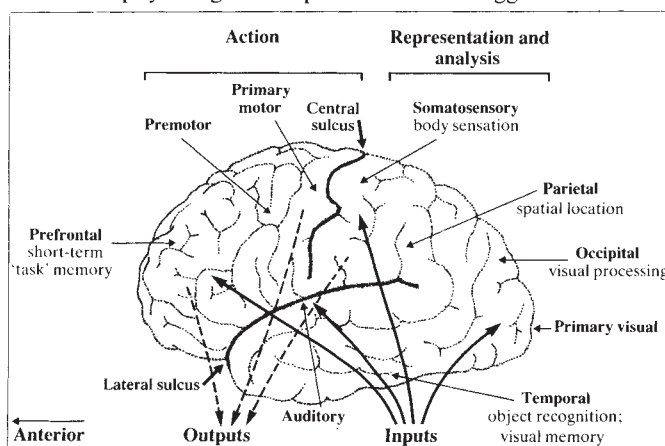
Another recent change has been the recognition that maps in the adult brain are much more labile than previously thought. Evidence for adult plasticity

comes from the intraspecific variation in maps and changes in the map with peripheral surgery⁹. Now, Michael Merzenich's group (University of California, San Francisco) has beautifully demonstrated that a map can change with use: the representation of a finger tip in the somatosensory cortex of the owl monkey expands after a period of intense stimulation, encroaching by up to 1 mm into the zones normally occupied by the rest of the finger and part of the hand. The grain of the map also becomes finer. The anatomical projection of the skin surface in the cortex is evidently more approximate and less discrete than implied by the physiological maps. Merzenich suggests

the inputs active at a given moment, but depends on the past history of activity in the neuron and the relative strengths of its momentary inputs. These include feedback from the motor system (Don Humphrey, Emory University, Atlanta) and from the catecholamine-containing neurons originating in the brain stem, which seem to be involved in arousal and attention (Steve Foote, University of California, San Diego; Floyd Bloom, Scripps Clinic). The response of a neuron to a specific stimulus will therefore be influenced by the behavioural state of the animal. Recording in anaesthetized animals is likely to reveal only the lowest common denominator of a neuron's full range of responses.

Although new techniques may provide new insights, more immediate progress is likely to come from asking the right questions with present methods (Charles Bruce, Yale University; Humphrey). It is only now being appreciated that the complexity of cortical function cannot be determined solely by building up from the responses of single units, and that the reductionist approach must be married with the 'top-down' analyses provided by ethology and psychophysics. Emilio Bizzi (MIT) stressed that this applies not only to sensory processing but also to the motor system, where the mechanics of the periphery determine what motor commands the nervous system needs to make.

It is also essential that the cortex is not considered in isolation. As well as the plethora of intracortical connections, there are extensive reciprocal connections between many cortical areas and the thalamus and basal ganglia, and indirectly with the cerebellum, brain stem and spinal cord. Deep understanding of how the cortex functions will be achieved only by setting the reductionist analysis in the context provided by the global operations of the whole nervous system and the behaviour of individual animals. □



The complexity of the anatomical terminology makes the cortical literature difficult to penetrate. This figure shows the main areas of the human neocortex, named according to their positions relative to the skull, and identifies their chief functions. The cortex is divided into posterior and anterior portions by two main folds, the central and lateral sulci. The rear half deals with the representation and analysis of information, with a separate area for each modality, such as vision, hearing and body sensation. Inputs of different modalities converge in the multisensory parietal and temporal lobes. In the front half, the motor areas are concerned with planning movements and determining how they will be executed; the prefrontal lobes provide short-term buffers for information needed to plan actions. The primary inputs come from the thalamus; outputs are to the basal ganglia and indirectly to the cerebellum, brain stem and spinal cord.

that response specificity at particular locations must be continually moulded by experience and the topography of the map maintained by dynamic processes throughout life. This type of plasticity may be a general feature of cortical maps and could underlie the acquisition of skills such as playing the piano or learning Braille.

The dynamic changes in maps seem likely to result from local interactions and modulation in the cortical circuits. David Prince (Stanford University) emphasized that the pyramidal cells, which generate the outputs from the maps, have numerous types of receptors for many transmitters and neuromodulators, which shape the activity of each neuron through a variety of effects on its many types of voltage-dependent ion channels. This means that the behaviour of a pyramidal cell cannot be predicted simply from its anatomical connections or by assessing

ganglia, and indirectly with the cerebellum, brain stem and spinal cord. Deep understanding of how the cortex functions will be achieved only by setting the reductionist analysis in the context provided by the global operations of the whole nervous system and the behaviour of individual animals. □

1. Shipp, S. & Zeki, S. *Nature* **315**, 322-325 (1985).
2. Hubel, D.H. & Livingstone, M.S. *Nature* **315**, 325-327 (1985).
3. De Yoe, E.A. & Van Essen, D.C. *Nature* **317**, 58-61 (1986).
4. Ungerleider, L.G. & Mishkin, M. in *Analysis of Visual Behavior* (eds Ingle, D.J., Goodale, M.A. & Mansfield, R.J.W.) (MIT Press, Cambridge, Massachusetts, 1982).
5. Rumelhart, D.E., Hinton, G.E. & Williams, R.J. *Nature* **323**, 533-536 (1986).
6. Gross, C.G., Rocha-Miranda, C.E. & Bender, D.B. *J. Neurophysiol.* **35**, 96-111 (1972).
7. Rolls, E.T., Perrett, D.I., Cann, A.W. & Wilson, F.A.W. *Brain* **105**, 611-646 (1982).
8. Perrett, D.I. et al. *Behav. Brain Res.* **16**, 153-170 (1985).
9. Merzenich, M.M. in *The Neural and Molecular Bases of Learning* (Dahlem Konferenzen LS38 (eds Changeux, J.-P. & Konishi, M.) 337-358 (Wiley, Chichester, 1987).

Jennifer Altman is an Assistant Editor of Nature.

IPG and DG, but there is disagreement about the structure of the DG, which may be derived from phosphatidylcholine rather than the precursor molecule for IPG (Farese).

Larner also showed that the rapid release of IPG (or an IPG-like compound) is abolished by pertussis toxin. Insulin-stimulated release of IPG is simultaneous with the release of alkaline phosphatase into the medium, and this effect is also abolished by pertussis toxin. The release of alkaline phosphatase is transient, and is important because this enzyme is anchored to the outside of the cell membrane through a glycosyl-phosphatidylinositol structure (Fig. 1). The insulin-induced release of heparan sulphate, a protein also anchored to the membrane by a similar structure, has recently been reported⁴. These experiments suggest a model for the release of IPG and DG. Insulin activates a phosphatidylinositol-specific phospholipase C. The action of the phospholipase cleaves off the DG moiety of the phosphatidylinositol, releasing the protein (alkaline phosphatase, heparan sulphate)-myoinositol-phosphate complex from the membrane. The released complex is then bound to a cell-surface transporter or receptor that recognizes and cleaves off the inositolphosphate-glycan and releases it intracellularly (Fig. 2).

Saltiel has reported that IPG can activate cyclic AMP phosphodiesterase and inhibit adenylate cyclase, but the full biological effect of this putative second messenger remains to be characterized. He has also recently reported the purification of an insulin-sensitive PI-PLC from liver plasma membranes⁵.

One of the confusing effects of insulin concerns serine phosphorylation. Most of the physiologically relevant target enzymes are dephosphorylated in response to insulin, but some are phosphorylated. Although most of the phosphorylation reactions seem to have no effect on biological activity, insulin appears to stimulate the activity of one or more uncharacterized serine kinases specific for ATP-citrate lyase or acetyl coenzyme A carboxylase (R. Denton, University of Bristol; see Fig. 2).

Can a unifying hypothesis be constructed from these results? The multiple effects of insulin occur in different time-frames, some taking place in a few minutes and others over several hours. It is possible, therefore, that different signalling systems mediate different effects. The phosphorylation of some enzymes in response to insulin may be caused by activation of serine/threonine kinases by the insulin receptor tyrosine kinase, whereas dephosphorylation may result from activation of phosphatases by mediators such as IPG. The diacylglycerol(s) could activate protein kinase

C. The activation of G_{ins} may also result in the inhibition of G_s dissociation by releasing β/γ subunits (Houslay) and hence in the inhibition of adenylate cyclase.

However, all of this is speculation. No evidence exists that proves conclusively any of the hypotheses I have described here. After decades of research, insulin continues to be the elusive pimpernel of endocrinology. □

Natural selection

It pays to be different

J. S. Jones and Paul H. Harvey

ON 25 June the Mall in London was filled with a diversely attired throng. Hundreds of guardsmen in bizarre but identical uniforms held back a crowd dressed for an English summer in a variety of brightly coloured raincoats. Through them, their royal leaders travelled in state to open a new parliament, one of whose tasks will, no doubt, be to ensure that any scientific research remaining in Britain is directed towards military or industrial ends. Those lucky enough to view this event from the tea room of the Royal Society could reflect on its pertinence to the theme of that day's scientific meeting*: frequency-dependent natural selection. In positive frequency dependence, any rare variant is at a disadvantage (as might be a French dragoon who mixed with the Guards at Waterloo), whereas in negative frequency dependence rare forms are favoured, a phenomenon familiar to the fashion designers who exploit the widespread desire to stand out in a crowd. The relevance of the Queen's badges of rank became obvious only later in the meeting.

Positive frequency dependence is important in understanding why distasteful species and their mimics evolve towards a common warning pattern and why species-specific mating signals are so strongly conserved. Negative frequency dependence is of more general interest as it is a powerful method of maintaining genetic diversity. If an allele is favoured as long as it remains rare, any new mutation will stay in a population until it reaches an optimum frequency above which it loses its advantage.

An association between the fitness and the frequency of particular species is at the centre of community ecology. J. Antonovics (Duke University) discussed experiments allowing the simultaneous analysis of frequency- and density-dependent selection. At low densities, the presence of plants of the same species has a much more deleterious effect on others than do equal numbers of plants of

1. Gawler, D. *et al.* *Nature* **327**, 229–232 (1987).
2. Saltiel, A. & Cuatrecasas, P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 5793–5797 (1986).
3. Low, M. *Biochem. J.* **244**, 1–13 (1987).
4. Ishihara, M., Fedarko, N.S. & Conrad, H.E. *J. biol. Chem.* **262**, 4708–4716 (1987).
5. Fox, J., Soliz, N.M. & Saltiel, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2663–2668 (1987).

Joe Espinal is at the Department of Clinical Biochemistry, University of Oxford, OX3 9DU, UK and is shortly to join Glaxo Inc., Research Triangle Park, North Carolina 27709, USA.

different species. Some of the best examples of genetic frequency dependence involve such interactions among inbred lines of plant, which prompted J.L. Harper (University College, Bangor) to accuse geneticists of re-inventing the ecological wheel, albeit without its spokes. Farmers have been doing population-genetics experiments for years. Wheat stands containing two cultivars (inbred and highly homozygous samples of a variable population) often give higher yields than can either component grown singly. This 'minority advantage' is intrinsically frequency dependent, as any deviation from the optimal proportion of each component reduces its fitness by, Antonovics calculates, as much as half. Such interactions may arise from differential use of resources. F. Christiansen (University of Aarhus) showed how genetic differences in resource use could maintain large amounts of polymorphism. But Antonovics suggests that in mixed patches of clones of the grass *Anthoxanthum odoratum* the minority advantage of rare genotypes arises not from resource differences but from an ability to escape from the pathogens which concentrate on the more common clones.

Coevolution of host and pathogen is strongly frequency-dependent. Any new mutation in a pathogen allowing it to circumvent the defence mechanisms of a host will rise in frequency until it exerts a selective pressure great enough to promote the spread of genes for new defences. Genetic mechanisms of protection in the host are exposed to similar selection. This relationship is seen in the 'boom-and-bust' cycle of sales of cultivars. In Britain a new variety of barley has a useful life of only 3–5 years before the spread of virulence genes in its fungal pathogens renders it useless. Models of host-parasite relationships (J. Barrett, University of Cambridge) have been quite successful in predicting such patterns in crop plants, but much less so in natural populations, as multiple infections of a single host and alternative breeding systems in host and parasite are difficult to

*Frequency Dependent Selection. Royal Society Discussion Meeting, 25–26 June 1987. Organized by B. C. Clarke, L. Partridge and A. Robertson.

analyse. The complexity of host-parasite evolution is emphasized by the computer simulations of J. Seger (University of Utah). In a model with two loci in the host determining a strain-specific defence against parasites, intermediate (but not extreme) rates of recombination produce cycles of coevolution. This could allow a parasite to act as an environment that fluctuates to favour the evolution of intermediate recombination rates in its host.

Rarity

Bacteria respond to infection with various restriction endonucleases, enzymes that attack phage DNA while avoiding their own. Phage counter this mechanism with enzymes that allow infecting DNA to mimic that of the bacterium. Any new restriction enzyme in the host or counter in the phage will be at an advantage when rare. The subdivision of many bacterial 'species' into genetically distinct clones may also reflect this interaction between host and parasite. Bacterial diversity might also be promoted by frequency-dependent resource specialization: cells resistant to antibiotics increase the fitness of sensitive cells once they become common as they detoxify enough antibiotic to allow the sensitives to grow (B. Levin, University of Massachusetts).

Interactions between predators and prey are analogous to those between parasite and host. J. Allen (University of Southampton) discussed experiments on artificial prey which show that predators concentrate on commoner morphs, perhaps because they form a preference for an abundant food item. New insights into the way in which predators obtain and process information from a visual environment show that cryptic patterning and frequency dependence may depend on very subtle cues (J. Endler, University of California, Santa Barbara). The predators themselves, in groups as different as monkeys and fish, possess polymorphism in the visual system that alters individual perception of prey. Rare morphs in cryptic and sparse prey gain a particularly strong frequency-dependent advantage. In dense populations, such morphs may be at a disadvantage as they stand out from the herd. Allen presented large numbers of dyed maggots to wild birds in a container heated by a candle: the warmer the can, the faster the movement of the maggots. The birds greatly prefer to attack prey that differ from the mass, particularly when they are moving quickly.

Individual recognition is also important in mate choice, and frequency dependence has often been invoked as a factor in mating success. As L. Partridge (University of Edinburgh) pointed out, much work on the 'rare male effect' in the fruit fly *Drosophila*, in which females are claimed to discriminate in favour of any genetically unusual male, is itself a can of

worms; it is bedevilled by statistical artefact and by irreproducible results. Disassortative mating does increase diversity at self-incompatibility loci in plants, and perhaps histocompatibility loci in rodents, as the cell-surface antigens of pollen and those secreted in rat urine are used for individual recognition by potential mates.

Mate choice in ladybirds is much better established than the non-specific preference for any rare male postulated for *Drosophila* (M. Majerus, Cambridge University). Some female *Adalia bipunctata* have a heritable preference for melanic males, the extent of which within a population changes in response to selection (although it has not been possible to select for a preference for non-melanic males). Although it is hard to see how fixed female preferences might arise at first, once they are established they set the fitness of males of the preferred genotype in a frequency-dependent way; when rare, preferred males will be accepted by relatively many females, but once they become common most females with the appropriate preference will already have mated. Models of this are surprisingly resilient to the chance of a female encountering a preferred male, and her readiness to accept a male of a less desirable genotype. However, the real world is — as usual — less tractable than is population-genetics theory, and various complications emerge as more is learned of the sex life of the ladybird. Females prefer dusky males even in the dark, and Glasgow females are choosier than those from Keele.

Aggression depends as much as does sex on individual identification. Many animals have displays that ensure that they fight only with rivals of the same status and avoid costly attacks on competitors that are almost certain to win a battle. Hierarchies such as these have a frequency-dependent component. An animal with a status identical to that of many others will have more potentially dangerous fights than will one with a standing sufficiently unusual to be seen only rarely as a potential opponent. Great tits, greenfinches and corn buntings carry 'badges of rank'; subtle cues in individual plumage which advertise their status and avoid unnecessary conflict. Natural selection might in this way produce a population that is genetically variable in behaviour or in pattern. D. G. C. Harper and J. Maynard Smith (University of Sussex) pointed out that such badges are evolutionarily stable only if the cost of fighting is high in relation to the value of the resource under contention, and if frauds (whose badges advertise a status beyond their capability) suffer high costs or are punished for their presumption. The scattered information now available does support these predictions, although it is clear that much more work is

necessary. For example, it seems that badge size is often correlated with brain size, heart size and fighting ability, so that, as the Queen has found, it is expensive to be a monarch even when relaxing at home. Current evolutionary models do not take this into account. Frequency-dependent selection acting on such quantitative characters can, however, maintain considerable genetic variation without the need for a separate mechanism of selection at each polymorphic locus (B. C. Clarke, University of Nottingham).

Population structure and environmental heterogeneity were discussed by several speakers. Frequency-dependent selection must occur if the world is patchy and organisms differ in fitness in the various patches. Patterning of clonal structure in plants leads to frequency-dependent interactions. This is most marked in response to pathogens, in which there can be very small-scale genetic differentiation in both fungus and host in a wild legume. In bacteria, too, colonies in soft agar (where subdivision is possible) show a greater frequency-dependent effect on the fitness of colicinogen genes (which release proteins that kill sensitive cells) than do those in chemostats. Rare morphs of palatable prey gain a further advantage if they are widely scattered, but warningly coloured and distasteful prey tend to cluster together. Mating preferences and kin selection are also influenced by population structure, as genetic subdivision alters the probability of relatedness among potential mates. In the same way, territorial animals are more likely to evolve polymorphisms in badges of status than are those that live in freely interbreeding groups.

Sceptics

At the end of the meeting it was clear that frequency dependence, like God, fails to convert its sceptics because it succeeds so well in explaining everything to its enthusiasts. Nowhere is this more clear than in the study of molecular evolution. An early flush of enthusiasm for this type of selection on protein polymorphisms has been succeeded by an eloquent silence. M. Nei (University of Texas) presented the alternative view: that of random change, in which the frequency of alleles for the great majority of molecular polymorphisms has no effect on their fitness. Neutral evolution is like original sin: everyone disapproves of it, but it certainly accounts for a lot. In contrast, frequency dependence is attractive to a variety of biologists it but has not yet succeeded in establishing itself as a unifying evolutionary principle. □

J. S. Jones is in the Department of Genetics and Biometry, 4 Stephenson Way, University College London, NW1 2HE and Paul H. Harvey is in the Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

Population biology

Three-year cycles of lemmings and Arctic geese explained

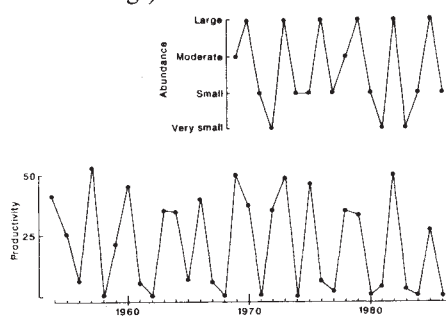
J. J. D. Greenwood

EVERYONE knows that many populations of lemmings vary enormously in size and that these variations follow fairly regular cycles. Arctic biologists are equally aware that these variations have profound effects on the predators of lemmings such as Arctic foxes: when lemming populations crash, the predators produce few, if any, young, despite turning to alternative forms of prey. Reproductive output is also variable in geese that breed in the Arctic, especially the dark-bellied brent goose *Branta b. bernicla*, in which the proportion of first-year birds in wintering populations is usually either less than 10 or greater than 30 per cent (see figure). Yet it was not until 1979 that Roselaar suggested¹ that the variations in the reproductive success of geese are linked to those in the populations of lemmings, through predators of lemmings turning to the eggs and young of ground-nesting birds in years when lemmings were not available. This suggestion seems to have been largely ignored until it was repeated recently by Ron Summers². The possibility is of more than academic interest: populations of brent geese have increased greatly in recent years and are causing agricultural damage, so they need to be managed — and successful management depends on understanding the population processes of the species.

The breeding productivity of the geese seems to show cycles with a periodicity of three years. Unfortunately, as Myrfyn Owen points out³, cycles of this periodicity are the most difficult to distinguish from purely random fluctuations using the usual statistical tests. Indeed, the number of turning points (the criterion of one standard test) expected in a random series is identical with the number that occur in a strict three-year cycle! Summers and Underhill⁴ have now subjected the brent goose data to more powerful tests, in which the lengths of runs of 'good' or 'poor' years, not just the number of runs, are taken into account: there is no doubt that the breeding output of the geese does, indeed, follow a three-year cycle, though with sufficient variation that it is difficult to predict output in advance.

Owen³ also points out that the original data presented by Summers are insufficient to show a formally significant correlation between reproductive output of brent geese and numbers of lemmings, but a slightly more extensive set of data shows a significant correlation⁴. Further, André Dhondt⁵ argues that the correlation to be

expected is not just a simple one of high output in good lemming years and poor output in lemming crashes. This is partly because the cycles are three years in length, not two, and partly because weather is likely also to affect the breeding output of the geese, as it affects that of many birds breeding in the Arctic. Dhondt suggests that the only definite prediction that can be made is that output of geese will be poor in years immediately after lemming peaks, as predators will then be common but will have no lemmings to feed on: the facts are that production was poor in 9 out of 10 such years. In peak years, when predator numbers may still be increasing (as they lag behind those of lemmings) and when there will be



Variation in abundance of lemmings (top) and productivity of dark-bellied brent geese (bottom) breeding in the Taimyr peninsula. Productivity is measured as the percentage of first-year birds in the succeeding winter population (data from ref. 4).

plenty of food to go round, geese would be expected to reproduce well, unless weather conditions were against them: production was, in fact, good in 6 out of 10 peak years, the four failures presumably being in years of poor weather. The situation in years preceding peaks will depend on how far the predator populations lag behind those of lemmings. If the lag is considerable then the increasing lemming populations will be more than enough for the still small populations of predators, and good reproduction of geese in those years would be expected unless the weather was poor: in fact, production was good in 9 out of 11 such years.

Weather has generally been considered to be an important cause of variation in the reproductive success of geese that breed in the Arctic. Summers and Underhill⁴ show that, although summer temperatures seem to have some effect on the productivity of dark-bellied brent geese, the effect is small compared with the apparent effect of lemming numbers. But

the most important meteorological information, the timing and distribution of snow cover in the early part of the breeding season, is not available and other measures, such as mean temperature and precipitation, have to be used as surrogates. Furthermore, Hugh Boyd⁶ points out that the correlations between the breeding production of Siberian brent geese and weather variables have themselves altered with time, as mean weather conditions, variability of weather and goose numbers have changed.

Boyd⁶ also notes that brent geese do not breed until they are three years old, so that a three-year cycle of breeding output is likely to persist for some time after a particularly good or particularly poor year. Thus, although weather does not itself vary on a three-year cycle, it could be responsible for triggering three-year cycles in goose productivity. But there is a final piece of evidence that implicates lemming predators: three species of wading birds that breed in the same areas as the brent geese also show marked fluctuations in reproductive output, which are correlated with those of the brent geese and with the lemming cycle. Because they mature more rapidly than the geese, the three-year cycles in the productivity of the wading birds cannot merely be intrinsic ones triggered by weather conditions. Also, their wintering areas, migration patterns and foods are different from those of the geese, so feeding conditions before the breeding season (which are known to be important in determining reproductive success in many Arctic birds) cannot be responsible for the three-year cycle.

A valuable test of the importance of lemming predators to the cycles of productivity of brent geese would be data for the goose population breeding on Svalbard (Spitzbergen), where there are no lemmings. Unfortunately, there are also no data. Indeed, the whole analysis of the population biology of brent geese is bedevilled by the lack of sufficiently detailed and sufficiently extensive data: the data-gathering has simply not kept pace with the growth of the population. Great advances could also be made if there was more collaboration between western European wildfowl biologists working on the birds in the winter, and those in the Soviet Union, who are able to study them on their breeding grounds. Until such collaboration occurs, the fascinating interplay of geese, predators, lemmings and weather cannot be fully understood. □

1. Roselaar, C.S. *Watervogels* 4, 202-210 (1979).

2. Summers, R.W. *Bird Study* 33, 105-108 (1986).

3. Owen, M. *Bird Study* 34, 147-149 (1987).

4. Summers, R.W. & Underhill, L.G. *Bird Study* 34, 161-171 (1987).

5. Dhondt, A.A. *Bird Study* 34, 151-154 (1987).

6. Boyd, H. *Bird Study* 34, 155-159 (1987).

J.J.D. Greenwood is in the Department of Biological Sciences, University of Dundee, Dundee DD1 4HN, UK.

Superconducting gems

THE compositions and structures of the two high-temperature superconductors $\text{La}_{1-x}(\text{Ba}, \text{Sr})_x\text{CuO}_4$ and $\text{YBa}_2\text{Cu}_3\text{O}_7$ are now well established; that their physical properties are less well known results largely from the fact that they are most easily synthesized as powders or ceramics. One therefore measures the properties of a collection of randomly oriented grains, which is unsatisfactory because the grain surfaces may behave differently from the bulk material, and because any anisotropy will be hidden by a measurement that averages over many different grain orientations. The successful growth of single crystals of $\text{YBa}_2\text{Cu}_3\text{O}_7$, as large as 4 mm, reported by Schneemeyer *et al.* on page 601 of this issue¹, is thus an important step.

Although several groups (see, for example, ref. 2) have succeeded in growing partially oriented thin films of $\text{YBa}_2\text{Cu}_3\text{O}_7$ (now familiarly termed YBCO), growth of macroscopic crystals from a melt has proved difficult. Because YBCO melts incongruently, the first crystals to form from a liquid of YBCO composition are not YBCO; liquids richer in copper and barium are needed to get YBCO to crystallize first. Schneemeyer *et al.* found that crystal size, morphology and superconducting properties depend strongly on the melt composition used.

Although the structures of the lanthanum- and yttrium-based superconductors differ, they are both layered materials containing planes of Cu–O. This structural anisotropy is reflected in the electronic properties of both compounds, as revealed by oriented measurements of critical current density and upper critical field. In $\text{La}_{1-x}\text{Ba}_x\text{CuO}_4$, the upper critical field is five times higher when the field is applied perpendicular to the Cu–O planes³; in YBCO, the critical current density is more than ten times higher parallel to the planes⁴. It is therefore expected that other properties of these materials will show similar degrees of anisotropy, which is why oriented measurements on single crystals will be so important. Although the critical current and critical field measurements can be carried out on crystals only 0.5 mm in size, many other types of measurement will be made easier (or even possible) by crystals of the size grown by Schneemeyer *et al.* Neutron scattering measurements, for example, require a large sample because neutrons interact so weakly with matter. And any surface measurement, such as optical reflectivity or X-ray photoelectron spectroscopy, requires careful preparation of the sample surface to remove impurities or oxidation; this will be easier with larger crystals. Laura Garwin

Fundamental forces

In pursuit of the fifth force

E. Iacopini

NEWTON's theory of gravity assumes that the source of the gravitational force is the inertial mass. As a consequence of this hypothesis and of Newton's second law, the acceleration of a body in the presence of only a gravitational field is independent of its mass (*g*-universality). The equivalence of inertial and gravitational mass is a principle in Einstein's theory of general relativity and in several other alternative theories of gravitation. It is for all these reasons that the *g*-universality has been tested many times with increasing precision. Using the results obtained with these tests, one can set limits also on the existence of new long-range forces, that will manifest themselves as an apparent *g*-universality violation, provided their source is not the inertial mass. The recent hypothesis of E. Fischbach *et al.* (*Phys. Rev. Lett.* **56**, 3–6; 1986) on the existence of a fifth force, popularly but misleadingly called antigravity, has given a new impulse to the field. Since then two experiments (Thieberger, *P. Phys. Rev. Lett.* **58**, 1066–1069; 1987 and Stubbs, C.W. *et al. Phys. Rev. Lett.* **58**, 1070–1073; 1987) have been completed that test the hypothesis, but come to contradictory conclusions.

The historical experiments by Eötvös in 1922 and Renner in 1935 set limits for $\Delta g/g$ at 3×10^{-9} and 7×10^{-10} , respectively. More recently, Dicke's and Braginskii's experiments pushed the accuracy to 10^{-11}

and 10^{-12} , respectively. These latter experiments, however, actually compare the accelerations of bodies in free-fall towards the Sun, that is, on a very different scale of distance. In 1986, Fischbach *et al.*, re-analysing the Eötvös data, claimed there was evidence of a fifth force, different from the four known fundamental interactions, repelling all bodies from the Earth in proportion to their baryonic charge or nuclear content. The result would be that samples with the same (inertial) mass, but different chemical composition would arrive at the ground at slightly different times in a Galileo-type experiment. According to Fischbach, the potential energy of two point-like masses m_1 and m_2 placed at a distance r is

$$V = - (Gm_1 m_2 / r) \times (1 - \alpha_{12} \exp(-r/\lambda))$$

where G is Newton's universal constant of gravitation; the first bracket is the usual form of Newton's law of gravity; and the second represents the short-range departure caused by the fifth force, with λ the range of the fifth force, and α_{12} is material dependent, proportional to the product of the two baryonic charges β_1 and β_2 , normalized to their respective masses.

Geophysical gravity determinations in mineshafts (Stacey, F.D. & Tuck, G.J. *Nature* **292**, 230–232; 1981) support the idea of a short-range departure from newtonian gravity with $\alpha_{12} \sim 10^{-2}$ and $\lambda \sim 1$ km. Fischbach's suggestion is that the

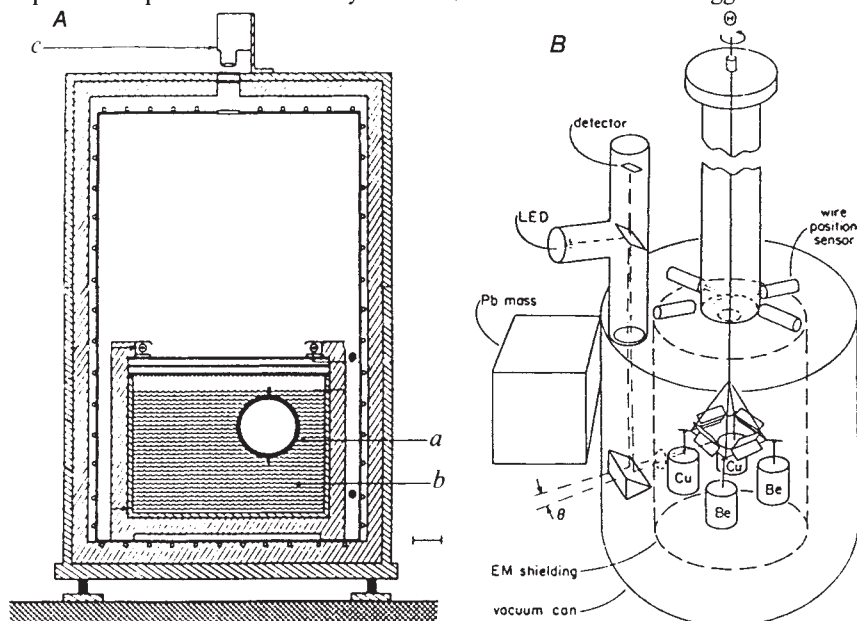


Fig. 1 A, Thieberger's accelerometer, in which the position of the hollow copper sphere (a) floating in distilled water (b) is viewed by the camera (c) once every hour. The whole apparatus is screened from electromagnetic effects and kept at constant temperature (scale bar, 10 cm). B, The torsional pendulum of Stubbs *et al.*, in which the differential force acting on the beryllium and copper cylinders is detected by the light-emitting diode (LED)–detector system as the pendulum rotates. The lead mass is to compensate for local gravity gradients.

1. Schneemeyer, L.F. *et al. Nature* **328**, 601–603 (1987).
2. Chaudhari, P., *et al. Phys. Rev. Lett.* **58**, 2684–2686 (1987).
3. Hidaka, Y. *et al. Jap. J. appl. Phys.* **26**, L377–L378 (1987).
4. Dinger, T.R. *et al. Phys. Rev. Lett.* **58**, 2687–2690 (1987).

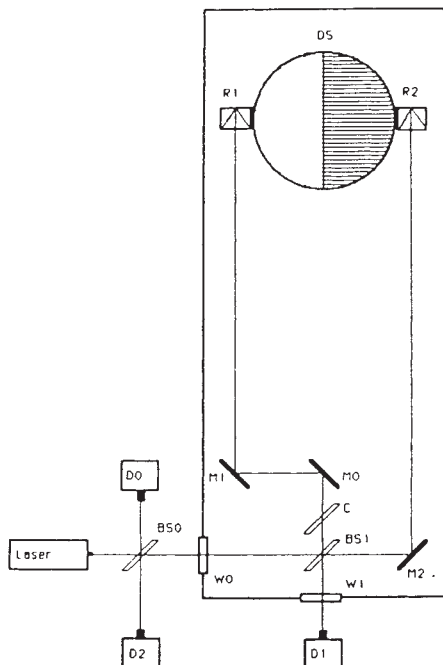


Fig. 2 The proposed Galilean experiment of Cavasinni *et al.* in which the fifth force would exert a torque on the disk (DS) made of beryllium and copper. The induced rotation would be detected by the Michelson interferometer. W0, W1, windows; BS0, BS1, beam splitters; M0–M2, mirrors; C, compensator; R1, R2, retroreflectors; D0–D2, photodetectors.

fifth force is mediated by the hyperphoton, a new particle that should also explain the anomalous properties of the sub-nuclear $K^0-\bar{K}^0$ system at high energies, but this hypothesis is open to serious criticism. It is more difficult to reject the correlation between $\Delta g/g$ and $\Delta\beta$ found in Fischbach's reanalysis of the Eötvös data; we actually need new experiments specifically designed to check the fifth-force hypothesis such as the new work by Thieberger and Stubbs.

Thieberger has developed a differential accelerometer in which he measures the horizontal motion of a well-balanced hollow copper sphere freely floating in pure water (Fig. 1A). If g -universality holds then the copper sphere, when balanced for water buoyancy, will remain balanced no matter what the gravitational configuration. Thieberger placed his accelerometer near the top edge of the Palisades cliff (New Jersey) and observed a horizontal drift of the sphere of 4.7 mm h^{-1} away from the cliff. The interpretation is that there is a horizontal component of force F resulting from the imbalance of the fifth force because of the cliff, that is stronger on copper than on water. This force repels the copper sphere from the cliff, and according to Stokes's formula has a value of $F \sim 4.2 \times 10^{-9}$ newtons, in good agreement with Stacey's geophysical data and with the value of 1.00171 for the ratio of the copper/water normalized baryonic charges. The experiment seems convincing and there are no doubts that the data support the presence of the force

F ; but does F arise from the fifth force or from something more conventional?

The question is justified by the contradictory result of Stubbs *et al.*. These authors have done an Eötvös-type experiment using a freely oscillating torsion pendulum, made with four cylinders, two of copper ($\beta=1.00112$) and two of beryllium ($\beta=0.99865$), having the same mass of 10.04 g and arranged in 4-fold symmetry (Fig. 1B). They measure the equilibrium position $\Theta(t)$ of the torsion pendulum placed in an evacuated vessel as it is rotated at constant angular velocity. The vessel is on the side of a hill at the University of Washington; because of the fifth force, there should be a time-modulated torque acting on the pendulum, inducing in $\Theta(t)$ a component at the vessel-rotation frequency. But they find no such component, in disagreement with Thieberger's experiment.

It is always difficult to draw definite conclusions from such delicate experiments. Thieberger's measurement should be repeated in different locations and with spheres of different composition and dimension. As for the experiment of Stubbs *et al.*, one would like to have more details on the determination of statistical error as well as on the systematic offset of 4 μ rad that they found.

Other experiments are in preparation. One proposal is to use the gravitational antenna (Amaldi, E. *et al.* *Nuovo Cim.* 9C 829–845; 1986) installed at CERN to detect gravitational waves emitted during

stellar collapses. If the fifth force exists, then the antenna would be excited by a dipole of baryonic charge, made from two hollow half-cylinders of iron and beryllium, rotating at the resonant frequency of the antenna (915 Hz). The shape and thickness are chosen to avoid ordinary gravitational effects on the antenna. A fifth-force signal corresponding to $\alpha=7.0 \times 10^{-3}$ and $\lambda=200$ m could be detected by integrating the antenna output for one day.

Another proposal (Cavasinni, V., Iacopini, E., Polacco, E. & Stefanini, G. *Phys. Lett.* 116A, 157–161; 1986) aims to check directly the g -universality with a Galileo-type experiment. A disk, made of two different materials, is allowed to drop in free fall and its angular motion is observed with a modified Michelson interferometer (Fig. 2). If the two half-disks fall with different accelerations, the disk will undergo an accelerated angular motion around its axis, proportional to Δg . The free-fall path foreseen is 5 m and the sensitivity is 10^{-10} in $\Delta g/g$, large enough to check Fischbach's hypothesis.

It is too early to draw any conclusions, but in a year or so there will be many experimental results, and we will know if there really is another fundamental force besides gravitation, electromagnetism, and the weak and strong interactions. $\square$

E. Iacopini is at the Scuola Normale Superiore, 56100 Pisa, Italy.

Estuarine chemistry

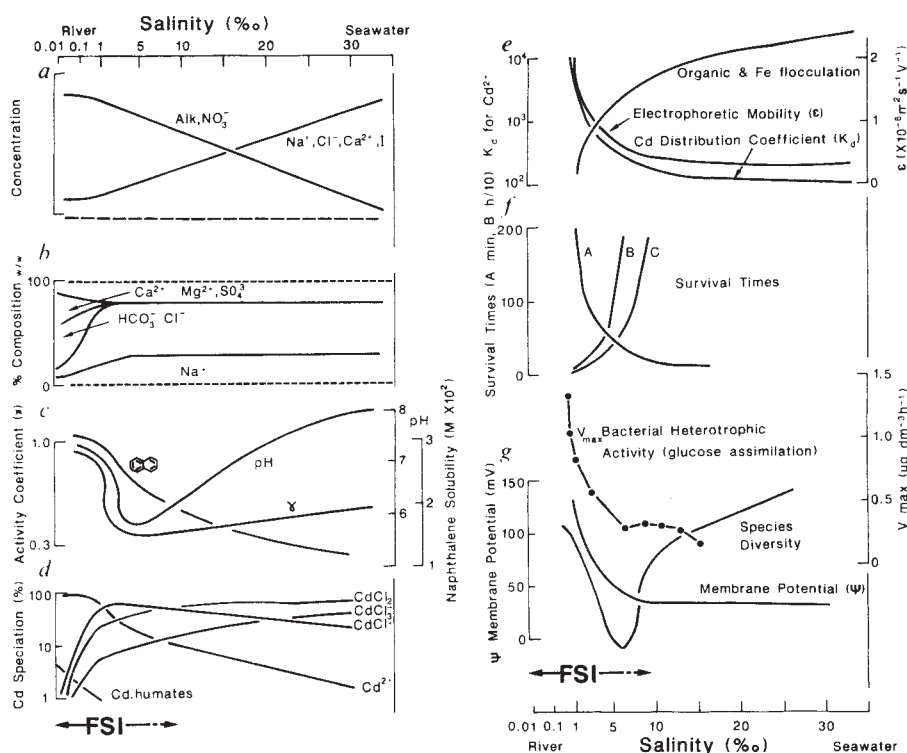
Organic films at the halocline

R. Fauzi C. Mantoura

WHEN rivers meet the ocean in the mixing zones of estuaries, a wide range of chemical reactions and biological responses are triggered, which are involved in the control of ocean chemistry¹ and the trophic and evolutionary segregation of terrestrial from marine organisms². Many of these chemical reactions have been successfully integrated into equilibrium thermodynamic³ and kinetic⁴ geochemical models which, supported by low-salinity data⁵, identify the fresh water-sea water interface (FSI) as an important site for triggering of biogeochemical processes in estuaries^{3,5-7}. Most of the estuaries that have been studied are fed by turbid rivers or are vertically well-mixed by tides. This has made it very difficult to deconvolute the axial gradients of salinity and dissolved chemicals in terms of the solution-state reactions within the FSI from the surface/solid-state reactions associated with advected and tidally suspended particles. In their paper on page 612 of this issue⁸, Žutić and Legović exploit the horizontal stratification of a salt-wedge estuary to

demonstrate, using elegant *in situ* microchemical techniques, the existence of stable biogenic surface-active films formed within a highly compressed FSI (30 parts per thousand, ‰, change within 20 cm) of the halocline. Not only have Žutić and Legović discovered a new family of biogeochemically important reactions at the FSI, but they have also identified the estuarine region in which heterogeneous particle reactions are spatially uncoupled from the unique solution-state chemistry of the FSI.

What is so special about the FSI? Although the estuarine distributions of many of the main dissolved constituents of rivers (Ca^{2+} , Mg^{2+} , alkalinity, dissolved organic carbon) and of sea water (Na^+ , Cl^- , Mg^{2+} , SO_4^{2-}) exhibit linear conservative dilution with respect to salinity¹ (see figure), both the aquatic thermodynamic properties and most other elements are highly variable or non-conservative in estuaries, particularly at the FSI. There are at least two fundamental reasons why the solution chemistry of the FSI is highly



Typical estuarine distributions of conservative (linearly diluted) and non-conservative chemical and biological constituents and properties in relation to estuarine salinity (S) with emphasis on the FSI. Note log scales for $S < 1\%$. In addition to the ions indicated in *a*, carbonate alkalinity (alk) and ionic strength (I) are also examples of conservative behaviour. Non-linear chemistry at the FSI is obtained for *b*, major-ion composition^{6,7}. *c*, Activity coefficients (γ) as derived from the Davies equation³ for triply charged ions (for example Fe^{3+}), pH and naphthalene solubility derived from the Setchenow equation. *d*, Speciation for dissolved cadmium including exponential variations in the mono-, di- and tri-chloro and humate-bound complexes³ at the FSI. *e*, Other FSI-reactive properties include organic flocculation⁹, particle-water distribution coefficient (K_d , ref. 12) for cadmium and electrophoretic mobility¹⁰ of estuarine particles. *f*, The biological perturbations include survival times for freshwater organisms such as *A*, the protozoan *Paramecium caudatum* and brackish-water organisms such as *B*, *Pygospio elegans* and *C*, *Amphiporus lactifloreus*. *g*, Variations in the membrane surface potential on *Amoeba proteus* and in total species diversity with salinity originated from Khlebovich¹¹ and Kinne². High bacterial heterotrophic activity expressed as the maximum assimilation rates for ^{14}C -glucose are also shown for the FSI of the Tamar Estuary.

non-linear with respect to salinity and is thus unusually reactive. First, all the changes in the compositional ratios of the major ions resulting from mixing river with sea water are complete within the first 1 ‰ of salinity. These are accompanied by massive concentration surges in the first 1–2 ‰ mixtures which have an overwhelming mass-action effect on chemical equilibria of the inflowing river water. This process is dramatically illustrated by the impact of Cl^- from sea water on estuarine speciation of dissolved cadmium (see figure). The dominance of the free Cd^{2+} species in river water is, within 3 ‰, dramatically reduced to less than 20 per cent by a sequential surge of three orders of magnitude in CdCl^+ , CdCl_2 and CdCl_3^- complexes. Humate-bound complexes of cadmium in river water are completely decomplexed by competition from Cl^- at a salinity of 1 ‰. Similar patterns of exponential changes in speciation at low salinities are predicted³ for other trace metals, such as Cu^{2+} , Zn^{2+} and Hg^{2+} .

A second, possibly more important, reason for the unusual chemistry of the FSI arises from the effect of increasing

ionic strength on ion activities. The Debye Huckel limiting law⁴ predicts that a linear increase in ionic strength (or salinity) results in an inverse hyperbolic decrease in ion-activity coefficients (γ , see figure), and hence a knock-on effect on reaction rates and stability constants that regulate processes such as hydrolysis, precipitation, complexation and colloidal stability^{4,9}. In the case of pH , the combined effects of different carbonate alkalinities and pH typical for river and sea water, together with the non-linear salinity-dependence of the first and second dissociation constants for carbonic acid, will give rise to a pH minimum at 0.5–3 ‰ in estuaries^{6,7}. Because pH controls the thermodynamic drive of many aquatic reactions¹, variations of 1.5–2.0 pH units observed at the FSI of several estuaries⁷ will have a profound impact on estuarine chemical processes such as dissolution of metal oxides and trace-metal adsorption of particulates. The solubilities and activities of unionized organics such as naphthalene (see figure) and gases in estuaries will also show greatest decreases at low salinities. The figure shows some examples of

the considerable changes in the chemical reactivity of particles during their passage through the FSI. These include electrostatic destabilization of amorphous $\text{Fe}(\text{OH})_3$ colloids, flocculation of organic macromolecules⁹, decrease in surface charge density¹⁰ and an up to 100-fold decrease in trace-metal binding strength on estuarine particles. As synergism can also exist between these and other FSI processes, it is clear that the addition of sea salt to a multicomponent equilibrium system of low ionic strength (river water) would result in non-linear chemical perturbations that are amplified within the low-salinity waters of the FSI.

In reviewing the salinity tolerance of aquatic organisms, Khlebovich¹¹, and later Kinne², identified a critical salinity boundary at 4–7 ‰ as a region of great physiological stress. In particular, osmoregulatory function, enzymic and ion-pump activities, membrane potential and survival times for both fresh-water and marine organisms (see figure) have all been shown to respond non-linearly with respect to salinity and decline steeply near the FSI, with a resultant dip in species diversity. We have suggested⁵ that mass mortality and degradation of fresh-water (and brackish water) phytoplankton at the FSI in the Tamar Estuary were responsible for the build-up of detrital organics and a coincident O_2 sag at the FSI. Žutić and Legović's results⁸ showing a strong enrichment of particulate and surface-active biogenic organics and pollutants may also render the FSI a site for pollutant incorporation into food chains.

Žutić and Legović's studies on the interfacial chemistry of the halocline at the FSI in the stratified Krka Estuary demonstrate an elegant field strategy for distinguishing the biogeochemical processes of the FSI from the 'fluidized-bed-reactions' of tidally stirred estuaries. Future studies on these globally important titrations at the FSI of estuaries could benefit enormously from the quantitatively more rigorous opportunities offered by similar salt-wedge structures that exist in many other microtidal estuaries, sea lochs and fjords.

1. Burton, J. D. & Liss, P. S. *Estuarine Chemistry* (Academic, London, 1976).
2. Kinne, O. *Marine Ecology* Vol. 1 Part 2 (Wiley Interscience, London, 1971).
3. Mantoura, R. F. C., Dixon, A. & Riley, J. P. *Est. Coastal Mar. Sci.* **6**, 387–408 (1978).
4. Stumm, W. & Morgan, J. J. *Aquatic Chemistry* 2nd Edn (Wiley, New York, 1981).
5. Morris, A. W. *et al.* *Nature* **274**, 678–680 (1978).
6. Mook, W. S. & Koene, B. K. S. *Est. Coastal Mar. Sci.* **3**, 325–336 (1975).
7. Morris, A. W. in *Environmental Biogeochemistry and Geomicrobiology* (ed. Krumbeyn, W. E.) Vol. 1, 179–183 (1978).
8. Žutić, V. & Legović, T. *Nature* **328**, 612–614 (1987).
9. Sholkovitz, E. R. *Geochim. cosmochim. Acta* **40**, 831–845 (1976).
10. Hunter, K. A. & Liss, P. S. *Nature* **282**, 823–825 (1979).
11. Khlebovich, V. V. *Mar. Biol.* **2**, 338–345 (1969).
12. Bale, A. J. thesis, Plymouth Polytechnic, 1987.

R. Fauzi C. Mantoura is at the Institute for Marine Environmental Research, Prospect Place, Plymouth, Devon PL1 3DH, UK.

The *ras* protein is not associated with exocytosis

SIR — In a recent News and Views article¹, Robert Burgoyne reported the suggestion that, because of the association of the protein p21 (the product of the *ras* oncogene) with pinocytosis² and its possible involvement with exocytosis, it could be the substrate for botulinum toxin type D.

Botulinum toxin type D blocks both cholinergic and adrenergic nerve transmission³. In the case of catecholamine release from cultured bovine adrenal medullary cells, where a rise in the intracellular Ca^{2+} normally triggers secretion, the toxin seems to act downstream of the Ca^{2+} transient at or near the site of exocytosis⁴. Because many toxins express their potencies by ADP-ribosylation of proteins^{5,6}, the findings that botulinum toxin type D ADP-ribosylates at least one membrane-bound protein of relative molecular mass 21,000 (21K) from bovine adrenal medulla homogenate⁷, and also from leaky adrenal medullary cells where the onset of ribosylation occurs in parallel with inhibition of exocytosis⁸, suggest not only a site of action of the toxin, but also the possibility that these ribosylated proteins are part of the machinery associated with exocytosis. Identification of these proteins in terms of the well-characterized proteins of other preparations⁹⁻¹¹ is an obvious next step. The evidence that we present here, however, clearly indicates that the *ras* product is not the protein.

The figure (A) shows that although botulinum toxin type D ribosylates proteins from bovine adrenal medulla homogenate, the toxin fails to ADP-ribosylate purified N- and Ha-*ras* proteins¹² under similar incubation conditions. The possibility that some endogenous factor from the homogenate is a necessary cofactor for ribosylation of the *ras* protein is unlikely

because incubation of the toxin-treated *ras* proteins with small amounts of homogenate gives only the pattern and extent of ribosylation corresponding to that seen with the homogenate alone, and thus no component of the ribosylation pattern can be attributed to the *ras* proteins. The toxin-induced ADP-ribosylated bands in both brain and medulla homogenates are not immunoprecipitated by incubation with a rabbit polyclonal antibody raised against purified p21^{N-ras}, whereas the iodinated *ras* protein itself is immunoprecipitated under identical incubation conditions (B in the figure). The failure to immunoprecipitate the ribosylated protein is unlikely to be a consequence of the ribosylation masking an immunoreactive site, as preincubation of homogenate with the antibody does not alter the subsequent toxin-induced ribosylation pattern. The suggestion reported by Burgoyne¹ is therefore not correct.

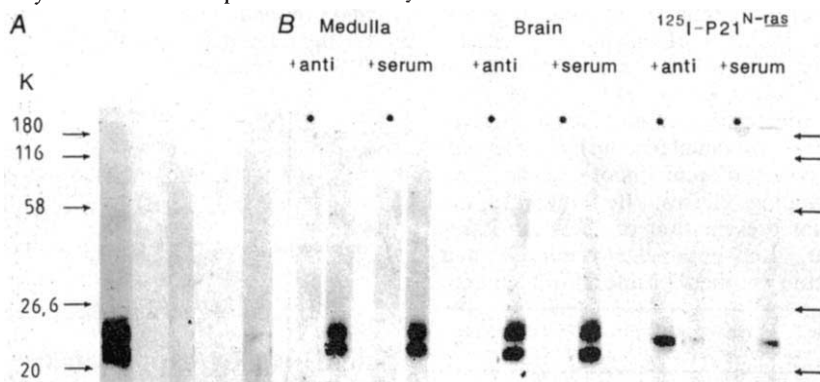
V. ADAM-VIZI
D. KNIGHT

MRC Secretary Mechanism Group,
King's College London,
Strand, London WC2R 2LS, UK

A. HALL

Institute of Cancer Research,
Chester Beatty Laboratories,
Fulham Road, London SW3 6JB, UK

1. Burgoyne, R.D. *Nature* **328**, 112-113 (1987).
2. Bar-Sagi, D. & Feramisco, J.R. *Science* **233**, 1061-1066 (1986).
3. Burgoyne, R.D. *Nature* **328**, 112-113 (1987).
4. Simpson, L.L. *Pharmac. Rev.* **33**, 155-188 (1981).
5. Knight, D.E., Tonge, D.A. & Baker, P.F. *Nature* **317**, 719-721 (1985).
6. Northup, G.K. et al. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6516-6520 (1980).
7. Katadi, T. & Ui, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3129-3133 (1982).
8. Ohashi, Y. & Narumiya, S. *J. biol. Chem.* **262**, 1430-1433 (1987).
9. Adam-Vizi, V. & Knight, D.E. *J. Physiol., Lond.* **394**, (1987).
10. Evans, T., Brown, M.L., Fraser, E.D. & Northup, G.K. *J. biol. Chem.* **261**, 7052-7059 (1986).
11. Kahn, R.A. & Gilman, A.G. *J. biol. Chem.* **261**, 7906-7911 (1986).
12. Salminen, A. & Novick, P.J. *Cell* **49**, 527-538 (1987).
13. Trahey, M. et al. *Molec. Cell Biol.* **7**, 541-543 (1987).



Autoradiograms of dried gels prepared using high-performance autoradiography films. A, The proteins p21^{Ha-ras} and p21^{N-ras} are not ribosylated by botulinum toxin type D; B, the proteins that are ADP-ribosylated by type D toxin do not react with polyclonal antibodies raised against *ras* proteins. In A, bovine adrenal medulla (a), purified p21^{N-ras} (b) and p21^{Ha-ras} (c) were incubated with activated toxin. In B, medulla (a-d) and rat brain cortex (e-h) were incubated with preimmune serum (+ serum) or with serum containing the antibody raised against p21^{N-ras} (+ anti) (a, c, e, g, i, k). Radioactive iodine-labelled p21^{N-ras} immunoprecipitated protein acted as a control (i-l). (For details of method, contact the authors.)

The other half of the global carbon dioxide problem

SIR—Your article “Half-truths make sense (almost)”¹ has missed half of the problem (and solution) by proposing only a supply-side solution. The rate of change of atmospheric carbon dioxide is associated with changes in its major sources and sinks, which are biotic, not industrial². The literature focuses on comparing quantities of carbon in different pools rather than fluxes between them, which are poorly known, except for combustion. This emphasizes large pools, which slowly exchange with the atmosphere (fossil fuels, soil humus, the deep ocean, limestones, volcanic gases) over smaller ones, which exchange much more rapidly (the tropical biota). Hence it artificially obscures the other key to the atmospheric CO₂ problem, which is to protect the metabolism of the tropical biota, the primary determinant of atmospheric composition.

Current land ‘development’ practices that ignore changes in community ecology have a large hidden price tag. They diminish³ the capacity of the biosphere to modulate atmospheric CO₂, which increases the magnitude and duration of climatic alteration^{4,5} beyond those forecast by conventional climate models. There are two real limiting options. First, destruction of biota and conversion of productive ecosystems to degraded secondary habitats increases combustion-derived atmospheric CO₂, which dissolves in the deep sea on a 10³ year timescale, and becomes incorporated into marine sedimentary carbonates and organic matter on a 10⁵ year timescale. Avoidance of the worst risks of the greenhouse effect will require global limits and national quotas on combustion¹. The second missing option is to conserve remaining undisturbed habitats that cycle carbon rapidly (like tropical rainforests and coral reefs) and to undertake pantropical replanting and fertilization to restore productivity to degraded areas.

Atmospheric CO₂ is held down by rapid recycling through the biota, and the ultimate fate of most added carbon is soil and sediment organic carbon, further enhancing soil productivity. To escape the greenhouse problem by renewable resource-based land management is much the cheapest option in the long run, and has many advantages. Solutions equitable to all countries demand cooperation. The price of fossil fuel should include its long-term environmental costs as an added ‘energy-growth’ tax, which transfers income from fuel burners to environmentally sound tropical development. The sooner such a course is adopted the greater the benefit and the lower the cost.

THOMAS J. GOREAU

Department of Geological Sciences,
University of Miami,
Coral Gables, Florida 33124, USA

1. Maddox, J. *Nature* **326**, 637 (1987).
2. Solomon, A.M. et al. in *Atmospheric Carbon Dioxide & the Global Carbon Cycle* (ed. Trabalka, J.R.) (USDOE, Washington, 1985).
3. Goreau, T.J. & de Mello, W.Z. in *Biogeochemistry of Tropical Rain Forests: Problems for Research* (eds Athie, D., Lovejoy, T.E. & Oyens, P. de M.) (WWF-CENA, Piracicaba, São Paulo, Brazil, 1985).
4. Goreau, T.J. & de Mello, W.Z. *Eos* **67**, 878 (1986).
5. Goreau, T.J. & de Mello, W.Z. *Ambio* (in the press).

Incubation time for AIDS

SIR—In “The sombre view of AIDS”¹, Rees suggests “a 15-year mean incubation and a high incidence of disease among carriers of the virus”. In view of the seriousness of the implications, we offer some critical comments on the manner in which these startling conclusions have been derived and presented.

Assuming a normal distribution, the mean incubation time was estimated from US data² on 145 transfusion cases of AIDS (acquired immune deficiency syndrome). The maximum observed incubation time was 7 years and only 18 cases had incubation times greater than 4 years. This does not, of course, imply that longer incubation times are rare, but simply reflects the low prevalence of human immunodeficiency virus (HIV) infection in the United States in the early years of the epidemic. But it does mean that, if the author's estimate of 15 years was correct, it would have been obtained from data belonging exclusively to the extreme left of the distribution (more than 1.5 standard deviations from the mean) and representing less than 1 per cent of the total cases among those infected in the period 1978–84. In the absence of any indication of the associated confidence interval, the result would thus have very little credibility, however correct the estimation procedure.

More specific criticisms can be made regarding the methodology used to determine the best-fitting normal distribution. First, no systematic search appears to have been carried out to investigate the effect on the chosen criterion of varying, independently, the two parameters (the mean, μ , and the standard deviation, σ) to be estimated. In fact several quite different combinations of μ and σ may yield essentially the same value of the chosen criterion (in this case the “total absolute difference” (TAD) rather than the more usual sum of squared differences). For instance, TAD = 14.67 with $\mu = 15$ and $\sigma = 5$, TAD = 14.47 with $\mu = 5$ and $\sigma = 2.14$, and TAD = 13.53 with $\mu = 10$ and $\sigma = 3.75$. Although not explicitly stated, the examples presented by Rees give the impression that, in searching to minimize the TAD, an arbitrary constraint was imposed by fixing the ratio μ/σ at 3. If this is the case, there is

Year of infection	Length of incubation period (years)						
	0–1	1–2	2–3	3–4	4–5	5–6	6–7
1978	0.07	0.12	0.20	0.32	0.50	0.73	1.06
1979	0.45	0.79	1.33	2.15	3.33	4.96	
1980	1.38	2.41	4.05	6.53	10.12		
1981	3.27	5.71	9.58	15.45			
1982	6.70	11.69	19.61				
1983	9.11	15.89					
1984	7.00						
Total	27.98	36.61	34.77	24.45	13.95	5.69	1.06

Expected cases ($\mu = 15$, $\sigma = 5$) classified according to length of incubation period
Total absolute deviation (TAD) = 31.8

absolutely no reason to suppose that the unconstrained minimum has been found.

Among several classical optimization methods available, the method of maximum likelihood³ has the advantage of being able to provide much needed estimates of confidence limits. This approach was used by Lui et al.⁴, who fitted a Weibull distribution to a subset (83 cases) of the same data set and estimated the mean at 4.5 or 5 years (according to whether multiple transfused cases were included or not) with 90 per cent confidence intervals of (2.6 years–14.2 years) or (2.3 years–19.6 years) respectively. More recently, Iversen and Engen⁵, also using a maximum likelihood method, have obtained similar estimates ($\mu = 5$ years, $\sigma = 1.6$ years) by fitting a normal distribution to the full data set. Regrettably, they did not give a confidence interval. Because they have used more data (150 cases), this could be expected to be smaller than that of Lui et al., but would nevertheless remain very wide.

The very large disparity between the above estimates and those of Rees can be explained. Peterman et al.² classified their data by year of transfusion and length of incubation period (in months). For reasons which are unclear, Rees has attempted to classify by year of transfusion and year of diagnosis (his Table 1). As the exact time of transfusion is not stated, the year of diagnosis cannot be ascertained precisely, thus leading to the use of 2-year classes (1978/79, 1979/80, . . .) which are not mutually exclusive. The TAD is evaluated from the ‘observed’ and ‘expected’ numbers of cases in these overlapping classes. If, however, the original presentation of the data is retained, for both the observed and expected numbers (Table 1), the column

Table 2 Comparison of calculated total absolute deviations for two pairs of parameter values

Parameter values	Total absolute deviation (TAD)	
	(a)	(b)
$\mu = 5$, $\sigma = 1.6$	35.6	23.2
$\mu = 15$, $\sigma = 5$	14.7	31.8

(a) cases classified by year of diagnosis (overlapping 2-year periods), as in Rees¹. (b) Cases classified by length of incubation period, as in Table 1.

totals represent the cases observed with specified incubation periods (0–1 years, 1–2 years, . . .), these classes being mutually exclusive. When the TAD is calculated in this way, quite different results are obtained. In particular, it now appears that the combination $\mu = 5$, $\sigma = 1.6$ does indeed provide a better fit to the data than $\mu = 15$, $\sigma = 5$ (Table 2).

Iversen et al. estimate (unfortunately without confidence limits) that 27 per cent of seropositive individuals will eventually develop AIDS. Rees's estimate of almost 100 per cent is based not only on his much more pessimistic estimates of the incubation time distribution but also on the assumed prevalence of HIV infection in the US and UK populations, estimates of which are known to be very imprecise.

In conclusion, while a mean incubation period of 15 years (or even more) cannot be ruled out, we believe that the best estimate at the present time remains at about 5 years, although with a very wide confidence interval.

DOMINIQUE COSTAGLIOLA
Unité de Recherches Biomathématiques
et Biostatistiques, INSERM,
Université Paris 7, 2 Place Jussieu,
75251 Paris Cedex 05, France

ANGELA M. DOWNS
WHO Collaborating Centre on AIDS,
IMET,
Hôpital Claude Bernard,
10 Avenue Porte d'Auvergniers,
75019 Paris, France

1. Rees, M. *Nature* **326**, 343–345 (1987).
2. Peterman, T. A. et al. *J. Am. med. Ass.* **254**, 2913–2917 (1985).
3. Rao, C. R. *Linear Statistical Inference and its Applications* 353–366 (Wiley, New York, 1965).
4. Lui, K.-J. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3051–3055 (1986).
5. Iversen, O.-J. & Engen, S. *J. Epid. Comm. Health* **41**, 55–58 (1986).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □

The case for the defence

Rudolf Peierls

Better a Shield Than a Sword: Perspectives on Defense and Technology.

By Edward Teller. *Free Press, New York/Collier Macmillan, London:1987.*
Pp.257. \$19.95, £17.25.

EDWARD Teller has often been called the father of the hydrogen bomb. Today he deserves the additional title of grandfather of Star Wars, since he is believed to have originated the idea which was publicized by President Reagan. Readers who expect, from the title of his new book, to find a full account of Teller's views on Star Wars, will be disappointed. It is a collection of 33 essays, of which only five are concerned with that topic, including the one that gives the title to the book.

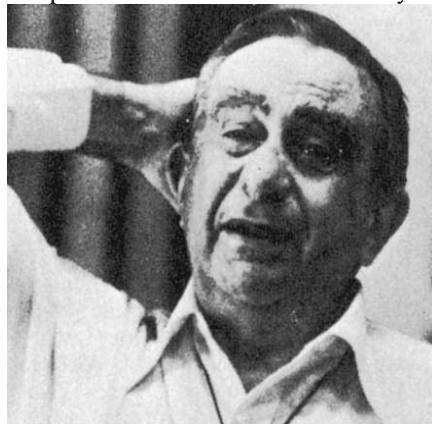
Thirty-nine pages is a short space in which to discuss a controversial subject in depth, and this is aggravated by the fact that the five essays were evidently written for different purposes (we are not told whether and where they were previously published), and there is much overlap and repetition. This section of the book therefore does not add much in the way of new arguments.

Teller makes much of the claim that the Soviet Union is well advanced in research on anti-missile defence, but does not give sources for this information. Presumably it comes from intelligence sources, but American intelligence has been known to exaggerate estimates of Soviet strength, presumably in order to support American defence projects.

The author is too good a physicist to claim that success in developing an effective defence is certain, and this is shown by his use of cautious phrases, for example: "Attempts to deploy a defense will be made only if there is a chance that defense will be less expensive than the compensatory deployment of aggressive weapons or the means to destroy the defenses" (p. 9); or "If we find real cost-effective ways for defense..." (p. 21). Yet in other places he sounds much more positive: "The genie that produced the sword of modern times can also produce the shield" (p. 22); or "Today there is increasing evidence that modern technology can produce successful defense measures" (p. 27). Perhaps the words "increasing evidence" are meant to convey a reservation, but this is not how the average reader will understand them. He is even more definite on page 31: "[the] development [of lasers] has produced the unexpected result of making defense feasible".

Many critics of the project have pointed out that the demands that the anti-missile defence make on computers far exceed in complexity anything so far achieved.

Teller disposes of this argument in rather a cavalier manner by claiming that the American telephone system is much more complex than the required defence management computer. It may be true that the 'phone system connects more subscribers than there are inputs and outputs in the defence system, but the computing functions in the latter are enormously more sophisticated. Moreover, the 'phone companies have to maintain an army of



Edward Teller — too good a physicist.

engineers to correct faults, whereas the defence computer would have to function reliably first time. All its parts could be tested separately, but the functioning of the whole can be rehearsed only in the real battle situation, when it is too late for corrections. There is no discussion of counter-measures, except for the use of decoys; none of the articles mentions the obvious response to laser beams against missiles of making the missiles spin.

I do not believe that these essays will convert any sceptic who doubts seriously whether the vast effort devoted to the Strategic Defense Initiative can make a contribution to anyone's security.

Several essays deal with the history of the atomic and hydrogen bombs, and it is interesting to have the recollections and reflections of someone so closely involved with both. During much of the wartime period at Los Alamos, Teller continued thinking about the hydrogen bomb, although this was not given any priority as it could make no contribution to the war. He explains the reason at one point (p. 53) by saying that Bethe asked him to take charge of the calculations on the implosion but "I refused. I knew there were others better at it. I also wanted to pursue more novel plans, including fusion".

This account agrees with my recollection; I took over that work instead, though I would not claim that I was better at it. However, in another essay (p. 71) we are told that Oppenheimer urged Teller to continue exploring the thermonuclear field, and he says "That was not easy advice for him to give or for me to take". In looking at recollections of events of some forty years ago one has to allow for uncertainties in human memory. Some little absent-mindedness is illustrated by his referring twice (pp. 45 and 205) to his stay at "City College" in London, when he means University College.

The essay "Seven Hours of Reminiscences" is a criticism of the BBC series on Oppenheimer (presumably written at about the time of the screening of that programme) which discusses Teller's view of Oppenheimer, whom he admires but finds hard to understand. It also outlines Teller's testimony at the Oppenheimer security hearings, where, in reply to the question whether he believed Oppenheimer to be a security risk, he explained that he often disagreed with him, and could not understand the basis of his actions: "In this very limited sense I would like to express a feeling that I would feel personally more secure if public matters rested in other hands".

In the essay he summarizes his "reluctant testimony": "I considered Oppenheimer loyal but because his actions appeared confused and complicated I would personally feel more secure if public matters rest in other hands." At the time, his testimony caused indignation among many of Teller's colleagues, who felt that his reply to the very clearly understood question whether there was a security risk was confused and complicated, and hard to understand.

The essay about the hydrogen bomb stresses in its title "the work of many people" and makes polite references to the contributions of Stanislaw Ulam and others, but leaves no doubt that the important invention was Teller's. "A Ray of Hope" comments enthusiastically on a plan for international control of atomic energy proposed by "a board of consultants to the State Department", although Teller suggests some modifications. This essay must puzzle the reader, until he discovers by looking at the list of copyright credits that it was written in 1946!

Other essays, too diverse to summarize, always present original and unorthodox views, which are interesting to read, whether one agrees with them, as with Teller's view that excessive secrecy does more harm than good, or disagrees, as with his ideas on the possibilities of Civil Defence in a nuclear war. □

Sir Rudolf Peierls is Emeritus Professor of Theoretical Physics in the University of Oxford, Oxford OX1 3RH, UK.

Structural survey

Janet M. Thornton

Introduction to Proteins and Protein Engineering. By Barry Robson and Jean Garnier. Elsevier: 1986. Pp. 699. Dfl. 475, \$190.

IRONICALLY it is the explosive growth in DNA sequence data and technology which has emphasized the urgent need for a deeper understanding of proteins and especially the relationship between sequence, structure and function. It has awakened in many experimental biochemists, immunologists and molecular biologists an interest in protein structure, and has generated the requirement for at least a minimal understanding as a basis for rational design of protein engineering experiments. In this context a new book on proteins and protein engineering is both timely and useful.

The text covers an enormous range of topics, from basic protein biochemistry, through gene manipulation to drug and vaccine design, and inevitably the emphasis reflects the interests of the authors. Robson and Garnier set out to produce a non-specialist text, comprehensible to a scientist of any discipline, which was written to be read rather than 'referred to'. This is a difficult task and underlies the book's strengths and weaknesses.

The book comprises three distinct parts. The first half is introductory, covering protein structure and function and gene manipulation. As well as the conventional descriptions of protein secondary and tertiary structure, there is the standard case study of a protein (lysozyme) and a chapter devoted to both metalloproteins and gene manipulation. Although these topics are now included in many texts they are well presented here, and some interesting historical perspectives are included. The central section describes current methods of structure prediction and constitutes almost one third of the book. Although this section begins at the general level, it becomes much more detailed and can only be described as specialist. Inevitably the authors draw heavily from their own research work, and it is sometimes difficult to extract the general principles from the detail. While it is appropriate for the scientist intent on structure prediction, it may be too difficult for the casual reader to follow. The final two chapters, forming the third part, concentrate on the background and methods involved in drug and vaccine design. These areas are at the forefront of current applied research and have immense importance in medicine. The authors have attempted to survey recent advances including, for example,

reference to interferons, antibody engineering, peptide vaccines and oncogenes. This is a good general introduction, although perhaps more emphasis could have been given to the results of recent experiments on the effects of single site mutations on protein and particularly enzyme action. Although many of these experiments are not directly relevant to drug or vaccine design, they illustrate the power of protein engineering and should, I feel, be included in an introductory text.

The authors have a relaxed, flowing style, which makes the book easy reading, although it does occasionally become verbose. The general standard of diagrams and presentation is excellent, with surprisingly few errors for a first edition. References are limited for most of the text, except for the section on energy

calculations, which has its own appendix.

For the most part, the authors have succeeded in their objective by making this a book to be read, rather than a reference text. Overall it provides an up-to-date description of proteins and the modern techniques of protein engineering and successfully demonstrates the potential range of applications of these techniques to modern medicine. Although too long for most undergraduate courses, it would be a useful addition to any research library and for any individual scientist who seriously wishes to explore for the first time the world of proteins and protein engineering. □

Janet M. Thornton is a lecturer in the crystallography department at Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK.

Tracts of wisdom

T. Scratcherd

Physiology of the Gastrointestinal Tract, 2nd Edn. Editor-in-chief Leonard R. Johnson. Raven: 1987. Two volumes, pp. 1,876. \$250.

THIS second edition of a work already established as a leading reference manual on gastrointestinal physiology is most welcome. There can be few books which can claim to be as comprehensive, with the possible exception of the American Physiological Society's *Handbook on the Alimentary Canal*, alas now out of date. It should be in all libraries, not only in those of universities and research institutes but also in the personal libraries of scientists and clinicians with a specialist interest in gastroenterology. In some ways the title is a misnomer: although most of the two volumes are concerned with physiology, pathophysiological topics, from peptic ulcers to parasites, are briefly considered.

This is a book for experts. In most chapters, evidence for mechanisms is precisely stated, citing the appropriate papers with a critical evaluation of the evidence presented. As a source of information it would be inappropriate for, say, a lecturer whose expertise is not in gastroenterology to use it to prepare a lecture. For example, should he wish to know the effect of the sympathetic nerves and adrenergic agonists on the pancreas, he would be faced with a plethora of apparently contradictory evidence, all correctly reported, and he would have to rely on the final conclusion that "No clear pattern emerges from the large number of studies on adrenergic regulation of exocrine pancreatic secretion": yet, clearly, adrenergic mechanisms can be demonstrated and common mechanisms

are present.

In some, but not in all chapters, there is a lack of historical perspective with little account taken of the development of ideas. Perhaps this is a price that has to be paid so that work published as recently as 1986 can be critically reviewed in a limited space. Hence this book does not make classical works, such as those by Babkin, Alvarez and Gregory, redundant.

The contributors have dealt with some of the most controversial topics in this field in an exemplary way: particularly useful are the detailed critical analyses of the current views of the inhibition of gastric secretion, the physiology of the 'enterogastrones' and the role of the cephalic phase in the control of exocrine pancreatic function.

In addition to the traditional subdivisions of motility, secretion, absorption and blood flow, the book covers functional morphology, immunology, intestinal flora, parasites and various aspects of pharmacology of the gut. Two rapidly expanding areas of knowledge which are in general poorly understood, the enteric nervous system and the regulatory peptides, are presented in a quite masterly fashion. Since the isolation and sequencing of gastrin and secretin, the gastrointestinal hormonal profile has proliferated to such an extent that most workers outside the field are left in bewilderment as to their function and how they fit into integrative control mechanisms. Dockray and Walsh have, to my mind, removed some of the barriers to understanding and made sense out of a puzzling morass.

This edition is completely revised and re-arranged. To those of us with presbyopia, the smaller print may be a handicap but the final product is worth this minor inconvenience. □

T. Scratcherd is Professor in the Department of Physiology, University of Sheffield, Sheffield S10 2TN, UK.

Deep friezes

Ann Sieveking

The Cave of Lascaux: The Final Photographic Record. By Mario Ruspoli. *Thames & Hudson, London/Harry N. Abrams, New York: 1987. Pp.208. £30, \$45.*

LASCAUX was closed to the public in April 1963, having, in the 19 years since it was opened, become one of the world's most famous prehistoric monuments. The huge numbers of visitors (there were 122,000 in 1962) caused a rise in temperature and humidity in the cave that allowed algae to grow on the walls and threatened to ruin the paintings. The French Government bought the cave, closed it and facilitated the creation of a copy, 'Lascaux II', for the benefit of tourists. In addition, they decided, in 1981, to make a full cinematic record of the original and Mario Ruspoli was commissioned to do this.

The book is his personal contribution to

this project: it is thoroughly researched and written for an interested general public. Although Ruspoli does not qualify his statements, and suppositions sometimes appear as facts, his appreciation of his subject, Palaeolithic art in general, Lascaux in particular and the way of life of its creators, is very well communicated. Certain aspects of the decoration, not apparent to an occasional visitor, were revealed by the conditions under which he worked. For example, his cameramen were restricted to using low lighting intermittently, and this sporadic and dim illumination exhibited the paintings as, perhaps, they were originally conceived. The friezes unfold before the viewer as he progresses, some figures recede, others emerge and some can be seen only from a succession of shifted positions. This is a situation that is familiar in Palaeolithic art; for example, the animals engraved on an antler shaft are curved around the convex surface so that they can be seen only in rotation — an aspect that still photographs and overall lighting cannot recreate.

Mud matters

Jeffrey Wilson

The Chemistry of Clays and Clay Minerals. Mineralogical Society Monograph No. 6. Edited by A.C.D. Newman. *Longman: 1987. Pp.480. £48. To be published in the United States by Wiley, \$110.*

CLAYS and clay minerals occur in geological environments as diverse as soils, sediments and hydrothermally altered rocks and are utilized for a remarkable variety of purposes, particularly in the ceramic, paper, chemical and oil industries. These materials therefore occupy an important niche at the interface between earth sciences and industry and are often the focus of interdisciplinary efforts by a whole spectrum of scientists. A consequence of the wide interest in clays is that information on their chemistry in relation to their industrially useful properties is scattered throughout the literature and can sometimes be difficult to find.

This book represents a major effort to bring together and to present authoritative information on two disparate but closely related themes, namely the chemical constitution of the silicate clays with their associated iron, aluminium and manganese oxides and hydroxides and the properties of these materials that are most useful to man, particularly from an industrial point of view. These properties include colloidal stability, ion exchange behaviour, thermal reactions, interactions with water and organic substances and catalytic properties. Inevitably, with a multi-author book of this kind covering so

many different aspects, there is a slight unevenness of treatment. For example, the approach adopted to describe the chemical constitution of clays is truly encyclopaedic and describes in detail not only those minerals that are clearly most important with regard to use, but also species that are at present probably of most direct concern to museum curators. In contrast, the cation exchange equilibrium of clays is dealt with in a mere ten pages of text and certainly some readers might have wished for more extensive coverage. Nevertheless, the overall balance between the chemical composition of clays on the one hand and the properties and behaviour of these materials on the other is about right, and the book will undoubtedly appeal to the international community of clay scientists and even to those with only a peripheral interest in clays.

The evolution of this book has not been without difficulty, as is frankly admitted by the editor in his preface. In fact, the project was initiated more than 15 years ago by the Clay Minerals Group of the Mineralogical Society and at one time, due to unforeseen circumstances, had almost foundered. Newman was appointed as editor and to supervise a rescue operation at a relatively late stage in the proceedings. It is a tribute to his patience and perseverance that the rescue was accomplished successfully and to his own scientific excellence that the final product is not just a bedraggled, out-of-date survivor, but a major contribution to clay science. □

Jeffrey Wilson is head of the Department of Mineral Soils at the Macaulay Land Use Research Institute, Aberdeen AB9 2QJ, UK.



Naive energy — possibly an illustration of a shamanic rite.

At Lascaux, the decorations exhibit an understanding of perspective in detail as in overall conception; for example, on many animals one leg is brought forward and the other placed in recession by erasing lines in the colour wash, and the topography of the cave is used as an inspiration for the layout of paintings and engravings. Here, as in all Palaeolithic caves, the decoration follows a strict formula in which certain animal species are associated, certain animals are depicted to suggest movement while others are static, and certain species are shown large or small. A marvellous variety of decoration is achieved within this format and the naivety of the figures gives them a particular energy. This is not a matter of anatomical depiction but is implicit in the size relation of the figures, the use of colour, the height of the friezes and the positions of the animals. Ruspoli's photographs communicate this alive and alert quality, although the plates are variable in the matter of colour. Lascaux is decorated in a warm spectrum and some of the photographs reproduced here are surprisingly cold.

Much of the technical information in the book is taken from *Lascaux Inconnu* (Arlette Leroi-Gourham *et al.*, 1979. Editions du CNRS) or contributed by scholars such as Brigitte and Giles Delluc, who also contributed to the earlier publication. In Ruspoli's book it is summarized and in English, which is a benefit to anyone who has not read the French original. Communication is a skill that few research scholars possess and Ruspoli's understanding of popular needs was an asset in writing an accessible and informative book. □

Ann Sieveking is an honorary research fellow at the London University Institute of Archaeology and author of *The Cave Artists* (Thames and Hudson, 1979).

Natural chemistry

Peter Karlson

A History of Biochemistry: Molecular Correlates of Biological Concepts. Comprehensive Biochemistry Volume 34A. By Pierre Laszlo. Elsevier: 1986. Pp.520. \$110, Dfl. 275.

THE subtitle of this book is somewhat misleading: "biological concepts" are phenomena such as hormonal regulation, nervous control of physiological function, interaction in embryonic development, physiological genetics and the immune response, to mention some examples. None of these fields is touched upon here. The subtitle was actually chosen by Florkin, the author of the volumes of *Comprehensive Biochemistry*. When Florkin died, Pierre Laszlo agreed to write this volume, and he felt there was "no good reason not to maintain [the subtitle]" (p. 485). But, as a chemist, he emphasizes the development of our knowledge of the chemical structures of molecules — protein molecules — rather than their biological function. Nevertheless, he tells an interesting story.

In part I, Laszlo discusses the structures of organic molecules in terms of their geometry, beginning with the ideas of the late eighteenth century, including the contributions of Kekulé (the benzene ring) and the concept of macromolecules (Staudinger, 1922); it also includes a discussion of mechanism versus vitalism over the centuries. Part II, "Proteins as Molecules", tells the story from the name 'protein', the contributions of Mulder and Berzelius in coining the name, to the identification of amino acids and the concept of peptides, and culminating in various proposals for protein structures. One chapter is devoted to Dorothy Wrinch and her cyclol theory, another to the contributions of Linus Pauling, and another to the sequence analysis of insulin by Fred Sanger. Part III discusses the history of our knowledge of proteins, using haemoglobin as an example, and part IV, "Enzymes as Proteins", contains a very interesting discussion of the old story of fermentation, the concept of enzymes as catalysts and the famous experiment of Buchner. It continues with binding of substrates to enzymes, the question of enzyme kinetics and finally X-ray studies on enzymes. Part V is a tentative comparative sketch of historical development in the haemoglobin and enzyme fields.

One criticism of the book is that it emphasizes the contribution of only a few scientists; many others are not mentioned at all. Consider, for example, the description of the haemoglobin story. The elucidation of the structure of haem as the prosthetic group of this pigment is

completely omitted, and the names of W. Küster and Hans Fischer are not mentioned. The elucidation of the three dimensional structure of haemoglobin by Perutz was greatly aided by the knowledge of the amino-acid sequence, elucidated by G. Braunitzer and his co-workers in 1959–61, again not mentioned. This is not the only case of imbalance and injustice. For example, Willstätter is described only as a man who, for a long time, pretended that enzymes are not proteins. It is true that this description of him, often repeated with his authority, held up the discovery of the true nature of enzymes. But one should give credit to Willstätter for his large amount of experimental work on various enzymes, including attempts at purification. It is mentioned briefly that Willstätter used saccharase as his favourite enzyme, the action of which could be measured by optical means, that is, the change in optical rotation. But from a

chemical point of view, this was a very poor choice.

On the other hand, many of the facts discussed in the text are well documented and illustrated by quotations in part from original work, in part from the Nobel lectures giving reflections of the Nobel laureates from the very beginning of their important work. The book also uses to a large extent quotations from other treatises on the history of biochemistry, providing the reader with a useful comparison of different views.

The reader will be delighted to find a lucidly written account of the ideas of various important scientists, including quotations from the original literature. But this book should not be relied upon entirely for the history of those aspects of biochemistry it describes. □

Peter Karlson is Professor Emeritus of Physiologische Chemie at the Philipps University, D-3550 Marburg, FRG.

Modern mammals

Gordon B. Corbet

Current Mammalogy, Vol. 1. Edited by Hugh H. Genoways. Plenum: 1987. Pp.519. \$75.

BEING SURROUNDED by a very diverse mammalian fauna that was relatively little studied until well into this century, it is understandable that many North American universities developed mammalogy as a discrete subject to a much greater extent than was normal in Europe. As a subject it performs two very useful compensatory roles: by encouraging consideration of the great diversity of mammals (over 4,000 species) it counters the tendency to use one or two 'types' to represent the class, whether in teaching or research; and it acts as a unifying agent, bringing together data and ideas from physiology, ethology, ecology, molecular biology and many other disciplines that employ mammals as subjects of study.

Mammalogy is already served by a number of more-or-less international journals devoted to research results and, to a lesser extent, review papers. However because of the dominant role of mammals in many disciplines, the volume of output of research far outstrips the capacity of existing review journals to effect a satisfactory synthesis. Symposium volumes go some way towards meeting this need but are usually on a very limited theme, comprise incomplete research reports or are produced in too much of a hurry to constitute adequate reviews of the subject. So there is indeed a niche for a series such as *Current Mammalogy* — provided of course that the articles meet a number of basic criteria, for example choice

of subject matter, thoroughness and authority.

Of the 13 chapters in this first volume (by 27 contributors, all but two from the United States) most are comprehensive, well-researched and refereed reviews of particular topics. Some chapters cover a broad spectrum, for example "Mammalian Evolution at the Cellular Level" and "Role of Chromosome Banding Patterns in Understanding Mammalian Evolution"; others are more specific, such as "The Behavior, Physiology and Anatomy of Lactation in the Pinnipedia". One chapter, "The Social Structure of Free-ranging Bottle-nose Dolphins", is in effect a research report, but it deals with a particularly thorough and long-term project that has been operating in Florida since 1970.

Ecological reviews are represented by a chapter entitled "Ants and Termites as Food", which tabulates data on 216 mammalian species, including such unexpected ones as the duiker antelope (*Cephalophus monticola*), in which 11 per cent dry weight of the diet consists of ants, and the mouse-deer (*Hyemoschus aquaticus*). A more wide-ranging subject is dealt with in "A Review of Density Dependence in Populations of Large Mammals", and applied ecology is covered by "Current Management Strategies for Commensal Rodents".

No predictions are made on the frequency with which subsequent volumes will appear, but if the high standards of this volume can be maintained the series will indeed provide a useful synthesis of a literature that is very widely scattered and highly fragmented. □

Gordon B. Corbet is in the Department of Central Services, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

How useful is Landsat monitoring?

B. Currey, A.S. Fraser and K.L. Bardsley

Early Landsat programmes were seen as the panacea for famine forecasting. But the images produced from the satellite are of doubtful value for monitoring rice production in southern Asia.

LANDSAT monitoring of crop production has been enthusiastically received as a breakthrough in famine prediction as well as in routine crop monitoring¹⁻⁷. However, previous crop-monitoring studies using Landsat have concentrated on corn, wheat and sorghum in the relatively dry climates of the American mid-west, the steppes of the Soviet Union and the Sahel of Africa^{8,9}. The potential problem of cloudiness in monitoring rice in southern Asia has been noted^{10,11}. But the use of Landsat continued because of the need for crop information in commodity forecasts for food security¹², the lack of emphasis on *usefulness* (as opposed to technical precision in the Landsat programme) and the assumed availability of useful Landsat images at particular dates in the rice crop calendar^{4,13,14}.

This article discusses the effect of cloudiness on the availability of useful Landsat images for critical dates in the cropping calendars at 58 rice nurseries in monsoon Asia. Three types of data were collected and analysed for each of the rice breeding and experimental stations of the

International Rice Testing Program (IRTP) and its cooperatives in monsoon Asia¹⁵.

Data were obtained from the US Earth Resources Observation System's Data Center on the image quality and on the percentage of cloud cover for every Landsat image of the 58 nurseries acquired during the 8.5 years between the launch of Landsat 1, on 23 July 1972, and 31 December 1980. For this period the availability for each rice nursery of band 5 images (as a proxy for bands 5 and 7 used for vegetation monitoring) was portrayed as in Fig. 1 by horizontal bars, perpendicular to the annual columns and located at the dates of acquisition.

The stippled horizontal bars indicate images judged to be 'inadequate' because of a 'poor quality' rating (<8 for band 5 on the National Aeronautics and Space Administration's subjective quality scale 1-8 for monochrome imagery). The 'adequate' images are portrayed as horizontal black bars. The length of a bar, measured in 10% intervals, indicates the percentage of a Landsat scene which is

cloud free. The longest bars represent images that were completely cloud free. It was assumed that the proportion of cloud cover over the 34,225 km² scene represented the cloud cover over the rice nursery. A conservatively high proportion of 40% cloud cover (that is, 60% of the image cloud free) was chosen to classify 'useful' images.

All IRTP nurseries grow early maturing (IR36) and/or medium maturing rice (IR38). Details of the phenophases (growth stages) for both were provided by the IRTP and its co-operatives¹⁵. The sequence of phenophases is shown as shaded tripartite sections in the vertical annual columns in Fig. 1. The phenophases were calculated from the planting, transplanting and flowering dates, assuming a month between flowering and harvesting unless otherwise stated. The vertical columns show the principal 'seedling', 'growing' and 'ripening' phases for both IR36 and IR38 for the final five years of the study, 1976-1980.

Additional data on regional cloud cover for 1979 and 1980 were provided by the

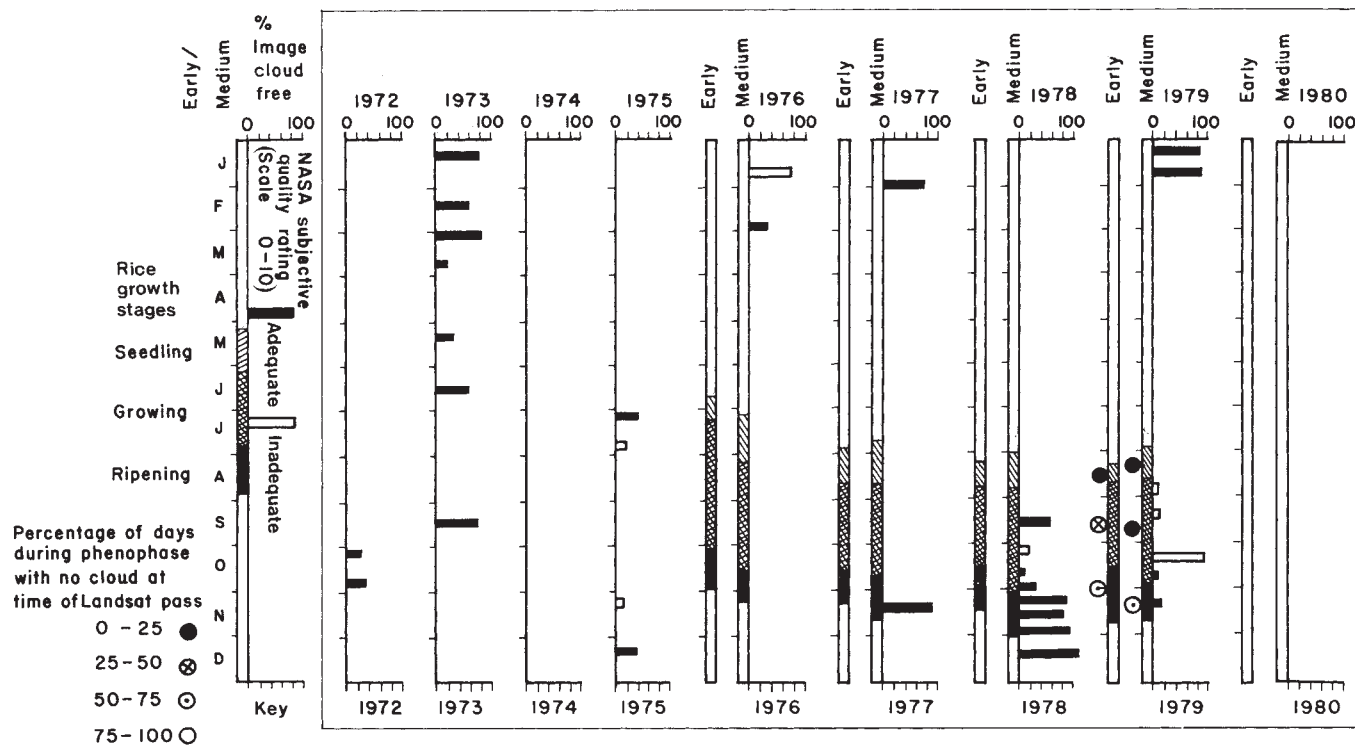


Fig.1 Overall availability and quality of Landsat images for Suphan Buri nursery, Thailand (14° 14' N, 100° 07' E). Also shown are the rice phenophase stages and cloud-cover data from the Geostationary Meteorological Satellite. See text for details. (Data courtesy IRTP, Australian Bureau of Meteorology and NASA.)

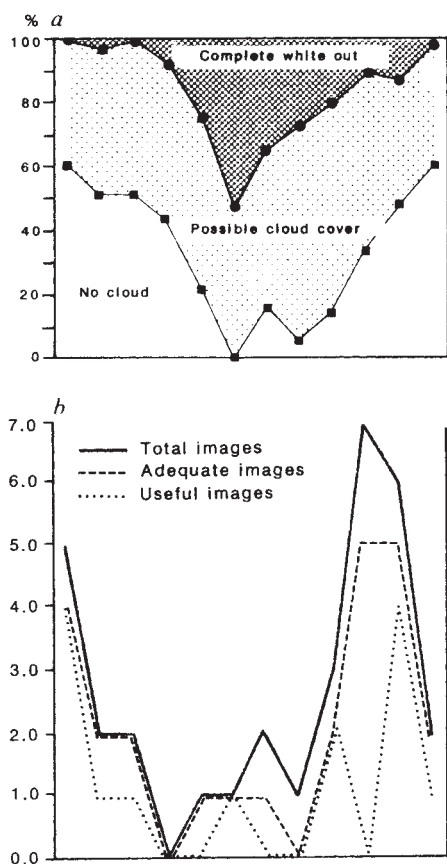


Fig. 2 *a*, Cloud cover 1979–1980 (as a percentage of days per month) over Suphan Buri nursery. *b*, Image availability (in number of images) for Suphan Buri nursery, 1973–1980.

Australian Bureau of Meteorology's interpretation of 'full Earth disk', visible and infrared images from the Geostationary Meteorological Satellite (GMS) positioned above the equator at 140°E. The GMS images are taken within a period 1.5 hr before or after the daily acquisition time for the Landsat image. The additional cloud data were collected only for the 45 nurseries east of 80° E longitude. The 13 nurseries west of 80° longitude were omitted because of the problems of curvature and parallax associated with the western extremity of the full-Earth-disk GMS images.

Because of the low resolution (≈ 5 km) of the GMS images, and the difficulties of distinguishing the light tone of cloud from that of snow or of other high albedo surfaces, the additional cloud information was classified into three categories only: (1) complete white-out, (2) possible cloud and (3) no cloud. On Fig. 1 the small circles adjacent to each rice phenophase indicate the percentage of days during the phenophase for which there was no cloud observed over the nursery at approximately the time when Landsat images were acquired.

Figure 1 shows image availability and cloudiness at Suphan Buri nursery in Thailand, which can be taken to be

representative of many nurseries. In this particular example, it should be noted that the total lack of images in 1974, a year of widespread famine in Asia, resulted from tape-recorder problems aboard the experimental satellite Landsat 1, and the saving of limited recorder capacity for coverage of the United States.

In the ripening periods of 1977 and 1978, long black bars indicate that cloud-free images of adequate quality were available. However, there was a distinct shortage of useful images for the seedling and growing phases in all years. Corroborative evidence is provided by the 1979 GMS cloud data which reveal that less than 25% of the days in early phenophases (seedling and growing) were definitely cloud free.

Effective regional interpretation of the data by aggregation of the phenophase and the image-availability data for all 58 nurseries, and by aggregation of the cloud-cover data, is constrained because of the location of many of the nurseries around the Bay of Bengal. Again taking Suphan Buri nursery as an example (Fig. 2*b*), there is a considerable difference between the peaks of total images in October, November and January and the trough in April, May and June for the seven-year period. The curve of adequate images has a similar shape. The shapes of these curves might be explained by the inactivity of the Landsat sensors at certain periods. Significantly, however, the troughs increase in the curve for the number of useful images (that is, those with 60% of an adequate image being cloud free). In general the curve of useful images (Fig. 2*b*) is congruent with the curve of possible cloud cover for 1979 and 1980 (Fig. 2*a*). At the Suphan Buri nursery the number of useful images was at a maximum in November at the end of the ripening phenophase (that is, at harvest time), but very few useful images were available before and during the seedling stage.

For the 58 nurseries (Fig. 3) there was a ninefold reduction in the availability of useful images from the dry to the wet season, a reduction attributable to increased cloud cover. The period of fewest days with no cloud and most days with complete white-out — June, July and August — is also the period of fewest useful images. The minimal Landsat image availability in July and August occurs at critical months for seedling and growing periods in many parts of the region.

In order to give a regional picture of the effect of cloudiness on Landsat's usefulness, the image-availability data for each of the 58 nurseries were mapped by month for the eight-year period (Fig. 3). In the map the radial bars represent the 12 months of the year plotted clockwise, with January at 12 o'clock. The full length of a bar indicates the total number of images

available for a month over the eight-year period. The black section indicates the number of useful images for that month. Many images were available for nurseries around the northern rim of the Bay of Bengal in India, Bangladesh and Burma.

There are, however, not only fewer adequate images available during the months of the Northern hemisphere summer monsoon (May–September), but there are virtually no useful images for nurseries such as Chinsura in West Bengal and Comilla in Bangladesh. The Landsat system acquires images, but cloudiness prohibits their use for monitoring rice during the monsoon months. The limitations of cloudiness associated with the summer monsoon are less obvious for nurseries in Sri Lanka. In the few Southern hemisphere nurseries of Muara, Bogor and Sukamandi in Indonesia, cloudiness associated with the maximum intensity of the Southern hemisphere summer monsoon in January prevents the acquisition of useful images. In January and February, cloudiness associated with the north-west monsoon over eastern China and the Gulf of Tonkin appears to reduce the availability of useful images. In Honiara in the Solomons there are no useful images available in any month of the year.

This investigation should continue with improved regional cloud-cover data from the higher-resolution meteorological satellites and ground observations of clouds in the rice nurseries themselves to determine the likelihood of obtaining useful images at particular times of year. Countries in monsoon Asia and famine-prone areas



Landsat image over the rice fields of Bangladesh at the confluence of the Brahmaputra and Ganges rivers.

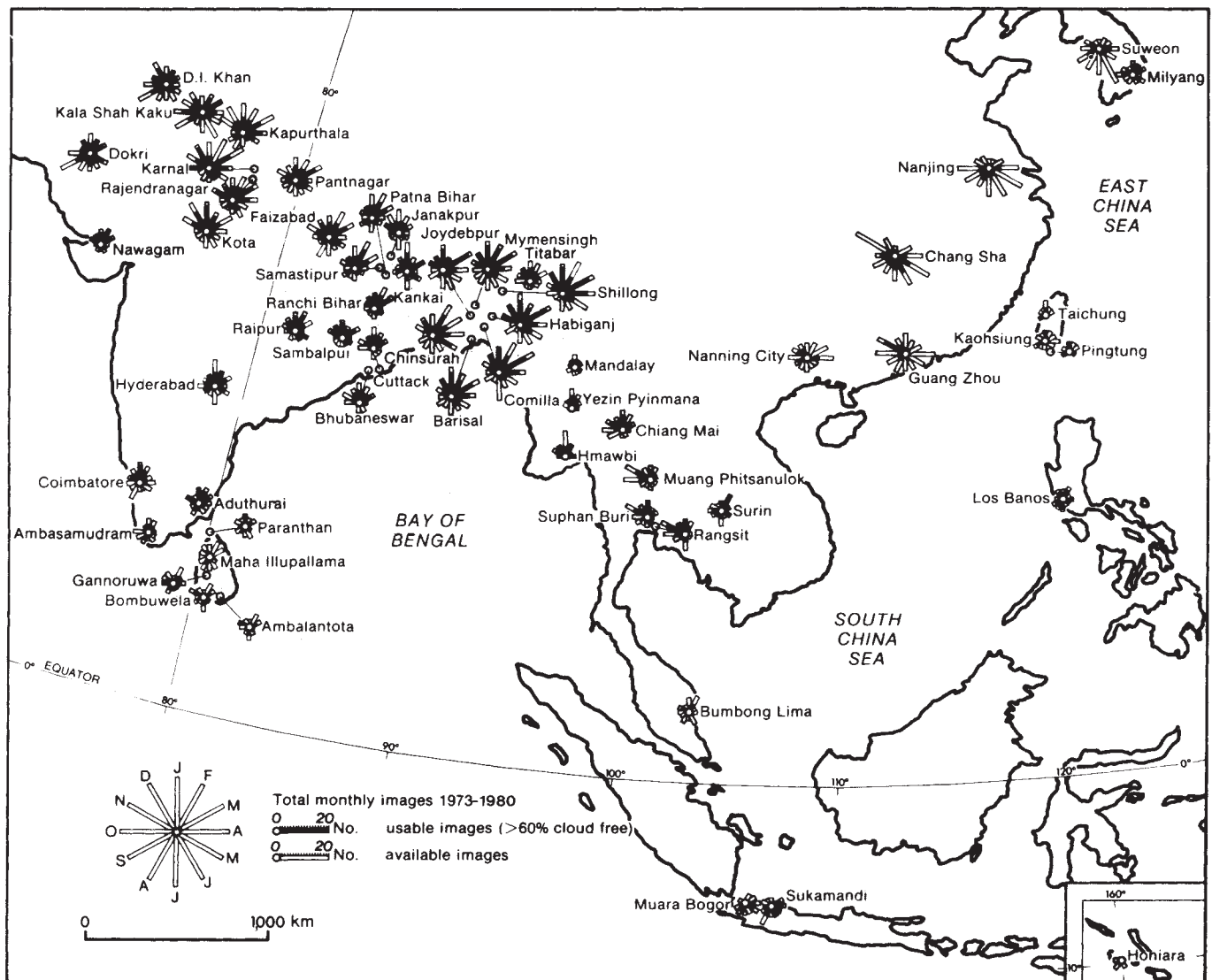


Fig. 3 Total images, by month, for all 58 IRTP nurseries during the period 1975–1980. See text for details.

elsewhere might have cloud-free times which are different from the Landsat pass time of 0930 hours¹ previously calculated as optimal for the United States.

Even as they stand, however, these results suggest that monitoring rice productivity by using only sequential Landsat images through all three phenophases with the minimum of one usable image per month⁴ is unlikely ever to be successful using the present timing of the Landsat pass over large areas of monsoon Asia. Landsat has a resolution of 80 m on the multi-spectral scanner and 30 m on the thematic mapper. Although the AVHRRs (Advanced Very High Resolution Radiometers) on NOAA 8 and 9 satellites have a lower resolution of 1.1 km, and technically permit regular coverage of areas in monsoon Asia four times daily to monitor regional vegetation indices and areas flooded, it appears that monitoring of crop productivity *per se* requires new forms of imagery.

After 1990, the benefits of radar for

penetrating cloud cover, seen from the image from the Shuttle Imaging Radar B, may provide useful imagery from the Canadian Radar Satellite. For famine forecasting, solutions must also be found to the complex links between crop yields and the community-level crisis of famine — even if there had been useful Landsat images available, the fields of Bangladesh would have appeared to be verdant in the famine years of 1974 and 1975.

1. Paul, C.K. & Mascarenhas, A.C. *Science* **214**, 139–145 (1981).
2. Paul, C.K. *Technol. Rev.* **82**, 19–29 (1979).
3. Morris, R.A. Letter. International Rice Research Institute, Los Banos, 11 Jan. (1981).
4. Malingreau, J.P. *The Wetland Rice Production System and Its Monitoring Using Remote Sensing in Indonesia* (University of California Press, Berkeley, 1980).
5. Haas, R.B. *Resources* June 11–12 (1983).
6. Cutler, P. *Disasters* **8**(1), 48–56 (1984).
7. Binder, A.K. *COSPAR: The Contribution of Space Observations to Global Food Information Systems* (Pergamon, New York, 1978).
8. MacDonald, R.B. & Hall, F.G. *Science* **208**, 670–679 (1980).
9. Adrien, P.M. & Baumgardner, M.F. *Science* **198**, 466–470 (1977).
10. Hossain, A. *Bangladesh National Programme of the Earth Resources Technology Satellite* (Planning Commission,

Dacca, 1974).

11. Committee on Remote Sensing for Development. *Remote Sensing From Space: Prospects for Developing Countries* (National Academy of Sciences, Washington, DC, 1977).
12. AGRISTARS. Annual Report FY-1981 (US Department of Agriculture/National Aeronautics and Space Administration, Washington, DC, 1981).
13. Lockheed Electronics Company Inc. *Cloud Cover Probability Study for the Large Area Crop Inventory Experiment (LACIE)* (National Aeronautics and Space Administration, Houston, 1974).
14. Smith, J.H. & Malin, J.T. *Technical Plan for Developing and Testing a Cloud Cover/Acquisition History Simulator* (AGRISTARS/National Aeronautics and Space Administration, Houston, 1981).
15. The International Rice Research Institute (1976–80) IRTP Final Reports of International Rice Yield Nurseries. (IRRI, Manila).
16. Sadler, J.C. & Harris, B.E. *The Mean Tropospheric Circulation and Cloudiness over Southeast Asia and Neighbouring Areas* (Hawaii Institute of Geophysics, Honolulu, 1970).
17. Polcyn, F.C. *Proceedings of the Seminar on the Application of Remote Sensing Techniques to Flood Hazard Assessment and to Flood Loss Prevention and Management* (United Nations, New York, 1985).

B. Currey is in the Human Resource Development Division, Winrock International Institute for Agricultural Development, PO Box 6083, Gulshan, Dhaka 12, Bangladesh. A.S. Fraser and K.L. Bardsley are in the Department of Geography, Flinders University, Bedford Park, South Australia 5042, Australia.

A high-resolution one-layer model of breaking planetary waves in the stratosphere

M. N. Jukes & M. E. McIntyre

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW, UK

Numerical integrations of a one-layer hemispheric model of the winter stratosphere have been carried out at very high resolution (of the order of a fraction of a degree of latitude). The numerical experiments simulate planetary-wave breaking and polar-vortex erosion under conditions far closer to the real winter stratosphere than those assumed in earlier, 'critical-layer' models. The model has implications for stratospheric photochemistry and the Antarctic ozone hole problem.

THE winter-time stratospheric circulation is dominated by a cold, cyclonic polar vortex and large-amplitude, planetary-scale disturbances on that vortex. The disturbances, whose amplitudes tend to be larger in the Northern than in the Southern Hemisphere winter, are usually thought of as Rossby waves originating in the troposphere, propagating upwards and then, typically, equatorwards as predicted by linearized wave theory^{1,2}. But two factors tend to make the waves strongly nonlinear by the time they reach the middle stratosphere, if not before. The first is the decreasing mass density encountered by the waves as they propagate upward, and the second is the weakening of the zonal winds as the waves propagate equatorward. Here we report the first results from a high-resolution numerical model study of the kind of nonlinear behaviour that might be involved. Recent observational clues about the likely behaviour and its implications for our understanding of the stratosphere have been discussed in a dozen or so publications (see, for example, refs 3–6 and their bibliographies); the present results give a clearer idea than before about flow features of smaller scale than can be resolved observationally, but which could well have significant effects on stratospheric dynamics and chemistry.

The breakdown of linear theory can be expected to lead to a complex flow, the nonlinearities generating many scales of motion and causing irreversible changes to the large-scale mean state⁶. Such flows should have some features in common with classical two-dimensional turbulence⁷. But the nonlinear interactions are highly non-local, both in physical and in wavenumber space, and the flow can be spatially very inhomogeneous, with strongly nonlinear regions adjacent to more wave-like regions—all of which indicates that classical homogeneous turbulence theories, while qualitatively suggestive, cannot be quantitatively applicable. Some insight into the way in which the different regions might interact has been gained through Rossby-wave nonlinear critical-layer theory (for example, ref. 8 and references therein); but that theory, which is restricted to small wave amplitudes and narrow nonlinear regions, is not directly applicable to the parameter regimes of greatest interest, with their large wave amplitudes and extensive nonlinear regions^{6,9}. In this situation, numerical modelling is the most powerful technique available.

Because horizontal resolution is expected to be at a premium, a logical first step is to use barotropic, or one-layer, modelling. There are two ways in which barotropic models capture some of the essential dynamics of the actual multi-layer, stably stratified, baroclinic system. First, both systems contain the basic dynamical mechanisms of Rossby-wave propagation and two-dimensional turbulence, reflecting the underlying analogy between baroclinic 'potential-vorticity dynamics' and barotropic 'vorticity dynamics'¹⁰. The baroclinic counterpart of two-dimensional turbulence is the layerwise horizontal or 'geostrophic' turbulence characteristic of strongly stratified flow^{7,11,12}. Second, major observed features of the flow in the real stratosphere often exhibit very deep vertical structures, especially in early winter (ref. 13; A. O'Neill, personal communication), so that barotropic modelling could also have some direct relevance.

Some thirty numerical experiments have been carried out, in which resolution and other conditions are varied and in which only the zonal harmonic wavenumber 1 is forced. In every case the results demonstrate the *in situ* nonlinear generation of a wide range of spatial scales, from wavenumber 2 down to scales well below observational resolution. Here we concentrate mainly on a case which illustrates, within a single high-resolution experiment, several typical aspects of the behaviour.

Numerical model

Equations of motion and experimental procedure. Some of the numerical experiments use the shallow-water equations and some, including the case to be described here, use the non-divergent barotropic vorticity equation in the form

$$DQ/Dt = \nu \nabla^6(Q - Q_0) \quad (1)$$

Here t is time and $D/Dt = \partial/\partial t + \mathbf{u} \cdot \nabla$, the material derivative; $\mathbf{u}$ is the velocity, taken to be horizontal and nondivergent. The quantity

$$Q = \zeta_a + F \quad (2)$$

where ζ_a is the absolute vorticity (ζ_a is equal to the Coriolis parameter plus the horizontal Laplacian $\nabla^2\psi$ of the streamfunction ψ). F is a prescribed, time-dependent, quasi-topographic forcing function, and Q_0 , a function of latitude only, is the initial profile of Q . The coefficient ν in the artificial, scale-selective dissipation term on the right of equation (1) is chosen to be a small as possible consistent with smooth numerical behaviour (see below).

The form of forcing chosen is the simplest for which the model in the limit $\nu \rightarrow 0$ has a materially conserved quantity, Q , whose spatial distribution can be compared with isentropic distributions of Rossby-Ertel potential vorticity in the real stratosphere. The idea is to imitate the effect of a planetary-scale Rossby wave incident from below upon a layer lying between two isentropic surfaces in the middle or upper stratosphere, as far as this can be done in a barotropic model.

The initial potential-vorticity profile Q_0 chosen for the experiment to be described here is shown as the heavy curve in Fig. 1a, and the corresponding zonal or azimuthal wind profile u_0 as the heavy curve in Fig. 1b. The coarsely dashed curve in Fig. 1a corresponds to the planetary vorticity profile, that is, the Q_0 profile for an unforced model stratosphere at relative rest. The finely dashed curves in Fig. 1a, b show another pair of initial profiles, to be referred to below. Both pairs give similar results. The numerical magnitudes are chosen to be broadly consistent with observed conditions in the middle and upper stratosphere, although the limitations of the model impose some compromises. The forcing function F takes the form of a stationary zonal

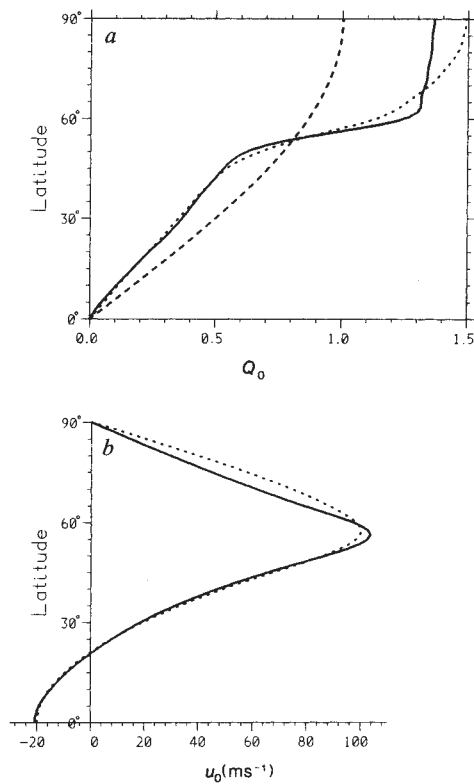


Fig. 1 *a*, Initial potential vorticity profiles Q_0 in units of the maximum planetary vorticity $2\Omega = 1.458 \times 10^{-4} \text{ s}^{-1}$. *b*, Initial zonal velocity profiles. The heavy, continuous curves were chosen for the experiment shown in Figs 3–5.

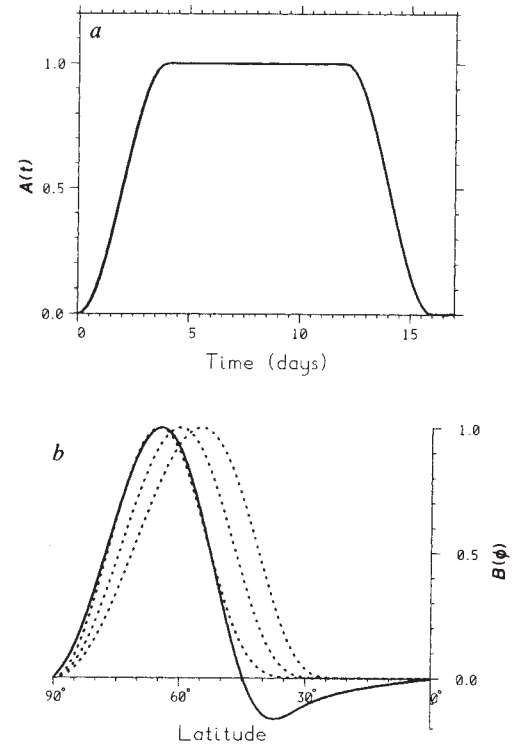


Fig. 2 The quasi-topographic forcing function. *a*, Temporal dependence $A(t)$; *b*, latitudinal dependences $B(\phi)$. The longitudinal dependence contains stationary wave 1 only. The heavy, continuous curves were chosen for the experiment shown in Figs 3–5.

harmonic wave 1 disturbance, in the form

$$F(\lambda, \phi, t) = F_0 A(t) B(\phi) \cos \lambda \quad (2)$$

where λ is longitude, ϕ latitude, and F_0 a constant. The function $A(t)$ used here is shown in Fig. 2*a*; it has continuous first derivatives and is made up of half-wave cosine and constant sections. Figure 2*b* shows several $B(\phi)$ profiles that were used in different experiments. The present experiment used the heavy curve; the results are not very profile-dependent. In the present experiment the forcing amplitude F_0 is taken to be 0.3 times the planetary vorticity at the pole. This produces disturbance amplitudes of the same order as those observed in Northern Hemisphere ‘minor warmings’.

Equation (1) is integrated numerically on a hemisphere, using a spherical harmonic expansion with triangular truncation Tn (truncation at a given total spherical harmonic wavenumber n) to give a homogeneous, isotropic representation. Resolutions up to T170 have been used. The present case used T159, which is finer than that used in ‘state of the art’ numerical weather prediction (T106), and equivalent to a mesh size well below 1° latitude. Hemispherical symmetry is assumed with Q , F and ψ taken as odd functions of latitude. Alias-free computation of the nonlinear terms is used¹⁴. At T159 each day of integration takes three minutes of processor time on a Cray-TS. The code is based on that written by Hoskins¹⁵, with modifications to take fuller advantage of the vectorization capability of the Cray; the numerical techniques used were originally developed by Bourke, Eliassen, Machenhauer, Orszag, Rasmussen and others¹⁵. A leapfrog time-stepping scheme is used with 300 steps per model day. A second order Runge-Kutta step is taken every model day to suppress the leapfrog splitting instability.

Model dissipation. The ∇^6 dissipation in equation (1) is used as a means of damping out enstrophy from the highest wavenumbers, whilst causing minimal disruption to other scales. At high

resolution it was found that simply putting the dissipation term into the leapfrog scheme did not work; for low values of ν the high wavenumbers became dominated by noise, while raising the value of ν made the scheme exponentially unstable. This problem did not occur at resolutions below T106. It was overcome by using the modified scheme

$$Q_{nm}^{k+1} - (Q_0)_{nm} = [Q_{nm}^{k-1} - (Q_0)_{nm} + 2\Delta t N_{nm}^k] \times \exp[-2\nu\Delta t\{n(n+1)\}^3] \quad (3)$$

which is stable for all values of ν . Here subscripts n , m refer to the total and zonal wavenumbers respectively, and superscript k to the time step. N is the spectral transform of the nonlinear term $-\nabla \cdot (\mathbf{u}Q) = -\mathbf{u} \cdot \nabla Q$. The value of ν used in this T159 experiment, $6.8 \times 10^{23} \text{ m}^6 \text{ s}^{-1}$, gives a dissipation timescale L^6/ν of 32 days for a length scale L equal to a degree of latitude (that is, for a total wavenumber of 57).

Results

Figures 3, 4 and 5 show daily potential vorticity and wind fields for days 0 to 17 of the experiment, omitting day 1 since it is almost indistinguishable from the initial state, day 0. The grey-scale shading represents values of the model potential vorticity Q , with higher values shown darker (except in Fig. 5*f*, see legend). The trajectories of a large number of material particles are confirmed independently, to check how accurately Q is conserved in the numerical integration. The locations of the material particles are marked with small crosses up to day 10; on day 0 they are densely spaced along two latitude circles near 40° N and 50° N (see Fig. 3 legend), giving the appearance of solid circles in Fig. 3*a*. After day 10 they are omitted for visual clarity and also because, by that stage, the model’s artificial dissipation is beginning to affect the smallest scales in the Q field. The close agreement between material and Q contours

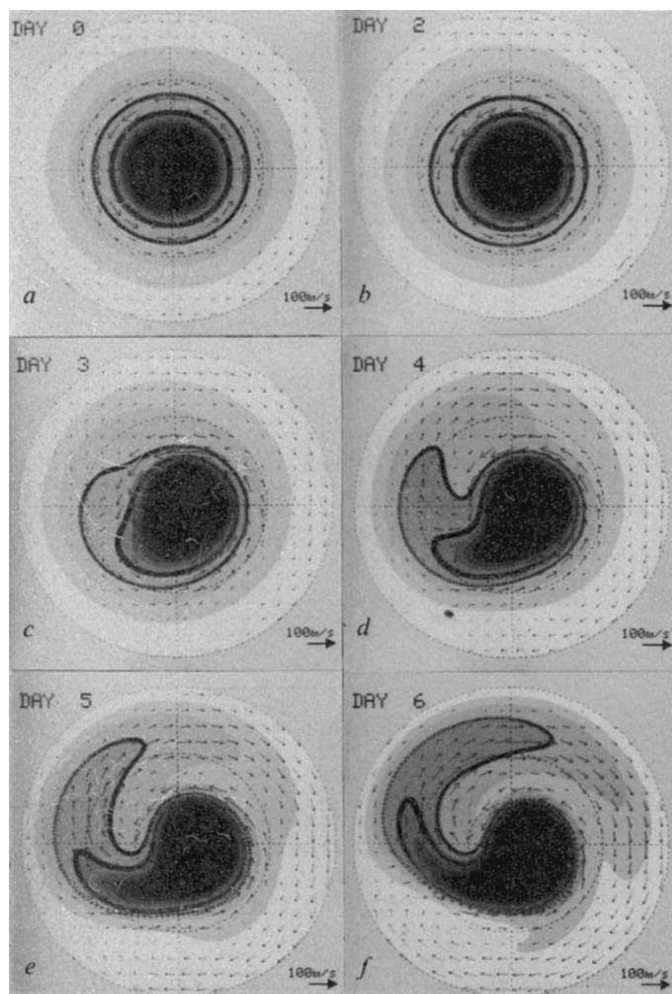


Fig. 3 The model potential vorticity Q and associated wind field at the times indicated. Dark shading corresponds to high Q , maximum value 1.36 times the planetary vorticity 2Ω , and light shading to low Q . The greyscale-density steps are at Q -intervals of 0.1515. Each arrow represents the wind velocity at the location of the centre of the arrow. The arrow length for a 100 ms^{-1} wind is shown at the bottom right. The maps are polar stereographic projections, with the 0° , 30° and 60° N latitude circles marked by dashed lines. The shapes of two material contours initially near 40° N and 50° N are computed independently of the Q evolution; the material particles are marked by small crosses. On day 0 these coincide with the $Q = 0.4545$ and 0.6060 contours.

before that time gives a stringent test of the numerical technique.

Three points are immediately apparent. First, the shapes of the contours show that linear Rossby-wave theory has ceased to be valid by day 4 and that wave 'breaking', involving irreversible deformation of the material contours shown, occurs thereafter. For a full discussion of the concepts of linear wave and breaking wave, as used here, see ref. 9. Second, the resulting 'surf zone' surrounding the main polar vortex develops ever smaller scales of motion as the integration proceeds, even though the forcing function F is a smooth, planetary-scale function involving zonal wave 1 only. Third, erosion of the main vortex occurs. For this purpose the main vortex is considered to be the material air mass marked by the central area of high Q -values. By day 17 this area is noticeably reduced in comparison with its initial value. Several authors (refs 4–6 and references therein) have argued that a similar erosion process may be significant for the dynamics and chemistry of the real stratosphere, including self-tuning resonance effects, although the finer-scale details of such a process cannot be resolved in present-day observational data.

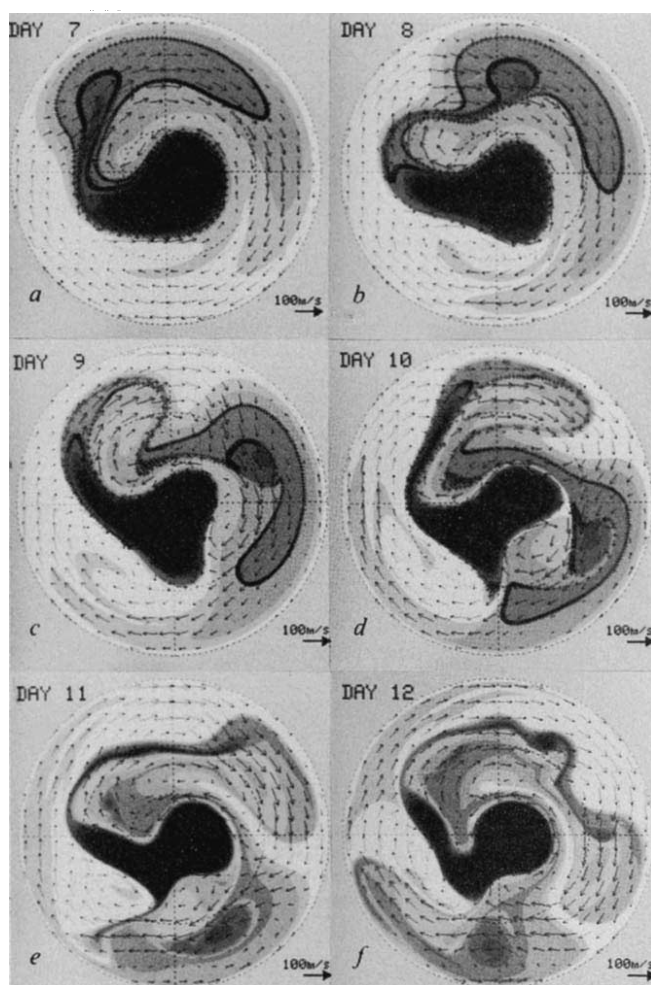


Fig. 4 Same as Fig. 3, but for days 7–12. Material particles are omitted after day 10.

As the numerical experiment proceeds, it is also noticeable that the Q gradients at the edge of the main vortex become very steep, well in excess of their initial values. This can be seen most clearly from the false-colour version of Fig. 5f reproduced on the cover of this issue. The tightness of the gradients that develop is limited by the model's artificial scale-selective dissipation, and it seems possible that even tighter gradients could occur at the edges of real stratospheric vortices, although the effects of baroclinicity and diabatic (radiative) damping would have to be considered for a full assessment. The same phenomenon has been seen in all the model experiments carried out so far, including some initialized with the finely dashed Q_0 profile in Fig. 1a. In every case it was found that gradients at the edge of the main vortex steepen to values limited by the model's artificial dissipation, over timescales of the order of a week or less. Such steepening, or 'interface-sharpening', is a well-known general characteristic of fluid-dynamical erosion-entrainment processes like those seen in laboratory experiments on mixing in density-stratified fluids¹⁶. There, vertical density stratification plays a role analogous to the horizontal Q stratification in the present numerical experiments.

Barotropic model experiments like the present ones are hardly likely to be realistic in all respects. One point on which the present experiments clearly differ from the real middle stratosphere is the way in which the surf zone penetrates too deeply into the tropics, producing an unrealistically large westward zonal-mean acceleration there (and, incidentally, giving a striking illustration of the connection between wave breaking and mean-flow changes^{4,9}). By contrast, real surf zones are more

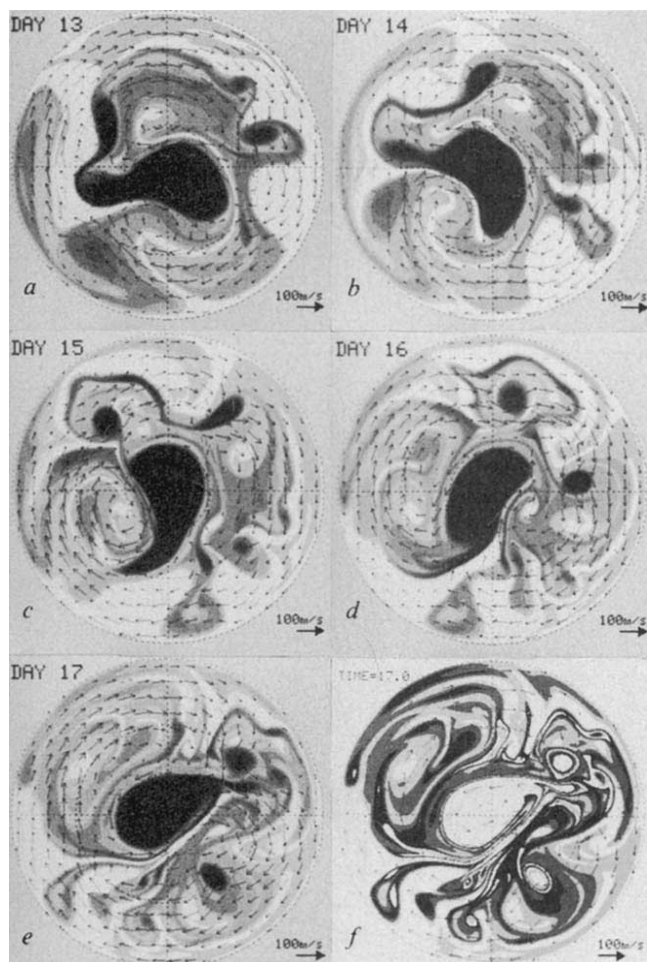


Fig. 5 Same as Fig. 4, but for days 13-17. The fields for day 17 are shown twice, the second time with a different shading scheme using the sequence white-grey-black, white-grey-black, white to cover the range in order of increasing Q values. In (f) the wind arrows are located with their tails at the relevant points.

confined to middle latitudes, the tropical middle stratosphere being comparatively little disturbed. This is shown by the far smaller zonal mean accelerations observed in the tropics and also, for instance, by the LIMS ozone maps shown in ref. 17. Ozone lifetimes should be long enough in the tropical middle stratosphere, around 10 days (for example, ref. 18, Fig. 1), to reveal the advective effects of disturbances like those in our experiment if such disturbances were present.

A second point is that the model's 'Aleutian anticyclone' is weaker, more elongated and less persistent than is often the case in the real stratosphere (day 5, top left quadrant). Both points might be related to the necessarily crude way in which the model disturbance is forced, as well as to the other limitations inherent in barotropic models, especially nondivergent ones, such as their well-known propensity to exaggerate the strength of the Rossby-wave propagation mechanism on the planetary scale. This in turn is related to the longer range of the two-dimensional inverse Laplacian operator giving ψ in terms of Q , as compared with the corresponding baroclinic potential-vorticity inversion operator¹⁰.

Despite these caveats some further features appearing in the model surf zone may be noted. During days 5-8, for instance, the wind field of the model Aleutian anticyclone is considerably strengthened by advection of low Q -values (light grey) from low to high latitudes. The same process has been found to be significant in an intermediate-resolution baroclinic model by O'Neill and Pope¹⁹, who argue for its importance in the real stratosphere as well. It is negligible, by contrast, in the critical-

layer models which provided some of the early illustrations of Rossby-wave beaking. Advection from low latitudes similarly leads to the formation of a weaker, second anticyclone near 20°W, 40°N on day 8, the end of the same period, taking 0° at the bottom of each picture. The resulting day-8 pattern would show a wave-2 component, among others, in a zonal harmonic analysis, even though the forcing function F contains no wave-2 component. The second anticyclone is a commonly occurring feature, and often a more prominent one, in the present series of experiments. It has been found to occur, typically near the end of the first week of integration, over a wide range of forcing and initial zonal velocity amplitudes, and in shallow-water as well as nondivergent model runs. An apparently similar *in situ* generation of wave 2 in the middle stratosphere has been remarked on both in observational data^{20,21}, and in experiments with a low-resolution model which allowed nonlinear interaction between zonal harmonic waves 1 and 2 only²¹. The present results support the tentative conclusion of ref. 21 that wave 2 may be generated as a result of local nonlinear dynamics rather than tropospheric forcing. Advective formation of a similar, but stronger, second anticyclone has also been remarked on in connection with the intermediate-resolution baroclinic simulations already mentioned¹⁹. Around day 9 of the present experiment, a third anticyclone is formed near 70°E, producing a more predominantly wave-3 pattern, again entirely by *in situ* nonlinearity.

Vortex rollup versus passive tracer behaviour

Another feature of interest appears on day 8. The high- Q air at the tip of the tongue produced by the initial wave-breaking event is rolling up into an approximately circular, cyclonic vortex at 170°E, 30°N. This vortex proves strong enough to resist immediate destruction by the large-scale deformation field. It is carried round in the tropical easterlies until day 14, without much change of form. On day 14, however, it enters a region of stronger large-scale deformation near 80°W and is quickly pulled out into a thin streak, as if it were a passive tracer.

On day 9 a second tongue of high Q is being drawn out from the vortex at 90°W, and on day 10 a long, narrow filament has been formed, extending further westwards on day 11. This can be compared to an idealized planar flow consisting of an infinite straight strip of high Q embedded in surroundings of lower Q . Such a flow is a steady solution of the barotropic vorticity equation, but it is unstable and tends to evolve into a string of rolled-up vortices connected by very thin filaments. A similar fate awaits the filament seen on day 10, though the instability is modified by the large-scale deformation field. By day 13 (Fig. 5a) vortex rollup has taken place, with two cyclonic vortices forming at about 110°E, 25°N and 140°E, 20°N. The cyclonic circulations associated with the instability and rollup do not show up very clearly in the wind field plotted in Fig. 5a because they are swamped by the strong westward flow.

The possibility that similar kinds of vortex rollup might occur in the real stratosphere, whether initiated by recognizable instabilities or in other ways, has been suggested⁶, and observational evidence for its actual occurrence in at least one case has been presented by Clough *et al.*¹³ That case was one in which the data quality could be checked particularly thoroughly. An isentropic map of potential vorticity from ref. 13 is reproduced here in Fig. 6, alongside a schematic drawing from ref. 6 which along with Fig. 5a suggests what this map might look like at higher resolution.

On the other hand, many of the thin filaments in the Q fields seen in Figs 4 and 5 show less tendency to exhibit vortex rollup. They behave much more like passive tracers. This appears to be related to the fact that the instability of a thin, straight strip of anomalous Q is easily suppressed by any large-scale deformation field which tends to elongate the strip; some relevant theory will be presented in a forthcoming paper. The strain rate required is only a small fraction of the Q contrast between the strip and

its surroundings (around a twentieth, depending on initial conditions). Also, a modest deformation field (strain rate $>0.15 \times Q$ contrast) can destroy a pre-existing vortex (for example ref. 22, p. 107). This is exactly what appears to happen to the cyclonic vortex near 80° W on day 14, as already noted.

As the flow continues to evolve, the surf zone becomes increasingly chaotic, the main vortex throwing off more vortices and filaments of cyclonic potential vorticity. By day 17 most of the region outside the vortex is filled with small-scale features, both vortices and filaments. Figure 5f shows this with a shading scheme giving greater contrast between adjacent Q values by using a multi-valued grey-scale (see Fig. 5f legend). The main vortex remains as a distinct material entity, albeit reduced in area by the erosion process, with very little transport into it.

Implications for stratospheric chemistry

Three aspects of the results are of interest in connection with stratospheric chemistry. First, they remind us of the likely importance of considering the main vortex as a material entity or isolated airmass, for dynamical as well as chemical purposes. We return to this question below. Second, the extent to which coherent vortices form in the surf zone is an indicator of possible differences between the small-scale behaviour of Q on one hand, and that of a dynamically passive tracer such as ozone or water vapour on the other (for example ref. 23). In the formulation of two-dimensional (height-latitude) models of stratospheric chemistry, for instance, the question arises whether the actual small-scale behaviour of Q and long-lived passive tracers can be parameterized, to some useful degree of approximation, via quasi-horizontal (isentropic) eddy diffusion coefficients (ref. 3, and references therein), and if so whether it is permissible, as a further approximation, to take the quasi-horizontal diffusion coefficient for Q to be equal to that for long-lived tracers at each height and latitude. The latter approximation would, if valid, provide a simple and extremely useful connection between the model's parameterized dynamics and its parameterized tracer transport²⁴. But the possibility of vortex rollup suggested both by the present experiment and by Clough *et al.*¹³ (Fig. 6b) warns that Q might not always diffuse as rapidly as a passive tracer. A quantitative study of this question following recent work by Pitari and Visconti²⁵ and Plumb and Mahlman²⁶ is now under way.

Third, the present results permit us to make a rough estimate of what might be called the 'mixdown time' in the surf zone, by which we mean the timescale for tracer features to be reduced in size to molecular-diffusive scales so that chemical reactions between previously separated constituents can begin²⁷. There is no implication here that the mixing need be either complete or spatially homogeneous; the typical timescales are nevertheless of interest in connection with the possible effects on chemical reactions having comparable or faster timescales.

To make a plausible extrapolation from the model to the real stratosphere, we make use of a standard idea from turbulence theory, that of 'random straining'²⁸⁻³⁰. This assumes that small-scale passive tracer behaviour is well modelled by subjecting the tracer to a deformation field of larger spatial scale and randomly varying strength and direction. The random-straining model predicts that, in the absence of molecular diffusion, the smallest horizontal scale l characterizing the tracer field will decrease exponentially as time goes on (in sharp contrast to the much slower, algebraic decrease produced by a steady shear or steady circular-vortex flow). Inspection of the smallest scales in Figs 4 and 5 (which are well resolved even on day 17, albeit somewhat affected by the model's artificial dissipation) suggests that l^{-1} has an e-folding time, τ_{dy} , of the order of a few days. This implies a time of the order of a month to reduce a tracer feature from the thousand-kilometre to the one-kilometre scale, and two months to reach a scale of metres in the horizontal. This process is slower than the corresponding process in ordinary three-dimensional turbulence, as measured against

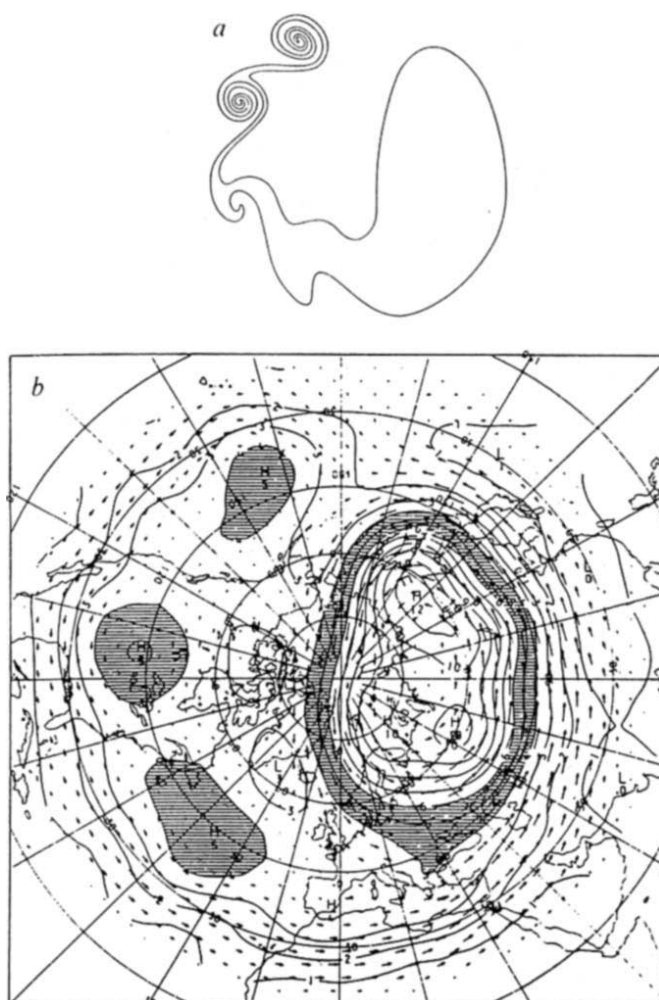


Fig. 6 a, Possible qualitative behaviour of a material contour originally near the edge of the main polar vortex, as speculated in ref. 6. b, Observational evidence for a similar phenomenon in the real stratosphere in a case where the data quality could be checked especially thoroughly, from ref. 13. Potential vorticity and wind arrows are shown for the 850 K isentropic surface (near 30-km altitude, or 10 mb) on 9 December 1981; observational resolution is too coarse to show any thin spiral filaments joining the vortices such as can be seen in Fig. 5a (upper right).

large-eddy timescales, because τ_{dy} is roughly scale-independent in two-dimensional 'turbulence'. By contrast, straining is enormously faster at small scales in three-dimensional turbulence because of the importance of three-dimensional vortex stretching.

In the real, baroclinic, molecularly diffusing stratosphere, we are interested in the vertical scale, h , associated with a given horizontal scale l , since it is vertical diffusion that becomes important first. Following the 'geostrophic turbulence' ideas of Charney⁷, we may reasonably assume that, as long as molecular diffusion is negligible, tracer structures having horizontal scale l will tend to have vertical scale $h \sim fl/N$, typically somewhat less than $10^{-2}l$. Here f is the Coriolis parameter, and N the buoyancy or Brunt-Väisälä frequency associated with the stable stratification. The same scaling $h \sim fl/N$ may be deduced directly from an extension of the random-straining idea to a baroclinic atmosphere, assuming a large-scale horizontal deformation field varying horizontally on scale $L \geq l$ and vertically on scale $H \sim fL/N \geq h$ (P.H. Haynes, personal communication). The essential point is that the associated vertical shear rapidly tilts any tracer structures that happen to have horizontal scale l but vertical scale $>h$, at such a rate that within one

horizontal deformation time they acquire slopes of order f/N and hence vertical scale h . Now molecular diffusion begins to become significant as soon as h becomes comparable to an appropriate diffusion scale. If we take the molecular diffusivity to be $\sim 10^{-3} \text{ m}^2 \text{ s}^{-1}$ (as for air molecules in the middle stratosphere), then molecular diffusion scales would be $\sim 10 \text{ m}$ for a time of 1 day, 30 m for 10 days, or 50 m for 30 days. An h of the order of tens of metres implies an l of the order of several kilometres, and hence, recalling the argument of the previous paragraph, a mixdown time of the order of a month or so.

As the heavier molecules will diffuse more slowly, and real surf zones may be less vigorous, and less unsteady, than the model one, we might expect mixdown times in a real wintertime surf zone to be longer, if anything, than a month, and even longer outside the surf zone. Some of the relevant chemical timescales are comparable to, or faster than this, suggesting the possibility of significant effects on the chemistry that might be missed by a simple diffusion parameterization. Nonlinearity in the chemical-kinetics equations makes the order in which different reagents are mixed, the rate of mixing, and the spatial correlations between the concentration fluctuations, all potentially significant.

The vertical mixing due to breaking internal gravity waves might modify the foregoing picture, especially when it comes to the detailed chemical modelling of small-scale effects. But the spatial intermittency believed to be characteristic of such processes implies that they are hardly likely to be capable of smoothing out the small-scale structure, as would an artificial, spatially homogeneous model eddy diffusivity.

Polar vortex as a material entity

The considerable ability of the model's main vortex to withstand large-scale disruption, despite its susceptibility to the removal of relatively small pieces from its edge, is characteristic of the behaviour of the model over quite a wide range of conditions. To some extent this robustness of the vortex is understandable in terms of the stability properties of two-dimensional circular vortices having monotonic radial Q gradients (see Fig. 1a). The stability theory goes back to Kelvin and Helmholtz, and modern developments are reviewed in refs 31 and 32. The stability of such vortices can be thought of in terms of the Rossby-wave propagation mechanism associated with the radial Q gradient. The associated restoring mechanism works most effectively on the largest spatial scales (for example, ref. 10, p. 920).

But stability is not the only consideration. In particular, it cannot explain the remarkable imperviousness of the main vortex even to small-scale incursions from the surrounding low- Q air. This 'one-sidedness' of the erosion process, whereby high- Q air is readily mixed into the surf zone but low- Q air less readily into the main vortex, is a typical characteristic of most of the model experiments, including those using the finely dashed initial profile in Fig. 1a. It is related *inter alia* to the existence of strong horizontal shear outside the edge of the high- Q region, and to the effect of that shear on the nonlinear behaviour of disturbances. The same one-sidedness has been strikingly illustrated in some recent work of Dritschel³³ concerning the behaviour of free disturbances in an idealized version of the same situation, in which the entire Q gradient is concentrated at a single material contour. In Dritschel's example the one-sidedness is complete; there is no transport into the model vortex, only erosion from it. In the present model experiments, it is possible to drive surf-zone material into the vortex if the forcing amplitude F_0 is increased sufficiently, but this seems to require very violent disturbances. In one such example (work in preparation), the vortex is nearly split in two by a large incursion of low- Q material, most of which is then thrown back into the surf zone.

That similar considerations might apply to the real winter stratosphere is suggested by the analogy, already alluded to, between the barotropic model's Q field and the baroclinic

stratosphere's potential vorticity (PV) field¹⁰. Isentropic distributions of Rossby-Ertel PV and wind speed observed in the winter stratosphere often resemble the model distributions shown in Fig. 1a, b in significant respects, despite the inability of a non-divergent barotropic model to match both distributions quantitatively. It therefore seems reasonable to hypothesize, on purely dynamical grounds, that the real stratospheric polar vortex, or rather whatever portion of it may have survived erosion, will often behave as a chemically isolated material entity even in the presence of substantial planetary-wave disturbances.

This hypothesis helps to explain observations concerning tracer behaviour both in the Arctic and in the Antarctic polar vortex, ranging from classical observations of ozone and volcanic aerosols (for example, refs 34–36) through to modern space and terrestrial-based observations³⁷. It seems reasonable to suppose, moreover, that the isolation will tend to be more complete in the case of the Antarctic vortex, because planetary-wave disturbances tend to be weaker in the Antarctic than in the Arctic. The strong vertical (entropy) stratification, and the strong radial (PV) 'stratification' at the edge of the vortex on each isentropic surface, may well succeed, especially in the Antarctic, in creating a kind of 'containment vessel' within which anomalous chemistry can take place as originally suggested by Farman *et al.*³⁸. Note that, as far as the Antarctic ozone hole problem is concerned, the chemically important region is thought to be the lower rather than the middle or upper stratosphere. But we believe that similar fluid-dynamical considerations will apply to the lower part of the vortex, the more so because of the longer radiative timescales there.

Further assessment of the vortex-isolation hypothesis will call for fully baroclinic simulations in which the vertical structure can be represented. If possible the horizontal resolution should approach those used here, and similarly the vertical resolution scaled by f/N . Such an assessment may well be critical for interpreting the results now becoming available from a number of observational programmes focused on the Antarctic problem. Numerical experiments on the United Kingdom's new Cray X-MP/48 are now being planned.

Whatever the precise degree of isolation of the vortex from its surroundings, the Lagrangian view of polar vortex behaviour seems certain to become increasingly important for our thinking about chemical modelling in general, as well as the ozone hole problem in particular. One implication is that the usefulness of low-order global models of stratospheric chemistry might well be enhanced if the model stratospheres are compared not with the real stratosphere averaged around latitude circles, but rather averaged around coarse-grain PV contours on isentropic surfaces so as to reflect the vortex structure. Some of the relevant ideas have been most recently discussed by Butchart and Remsberg⁵.

Support from the UK Meteorological Office, NERC, the Nuffield Foundation, the US Office of Naval Research the University of Cambridge Computing Service, and the University of London Computer Centre is gratefully acknowledged. We thank J. Austin, D. G. Dritschel, J. C. Farman, P. H. Haynes, R. L. Jones, P. A. Newman, A. O'Neill, V. D. Pope, J. A. Pyle, R. B. Rood, M. L. Salby, T. G. Shepherd and A. F. Tuck for helpful comments or correspondence, and B. J. Hoskins and I. N. James for kindly providing the original model code and for valuable advice on its use.

Received 5 April; accepted 1 July 1987.

1. Charney, J. G. & Drazin, P. G. *J. geophys. Res.* **66**, 83–109 (1961).
2. Matsuno, T. *J. atmos. Sci.* **27**, 871–883 (1970).
3. World Meteorological Organization *Atmospheric ozone 1985: Assessment of our Understanding of the Processes Controlling its Present Distribution and Change* (World Meteorological Organization, Global Ozone Research and Monitoring Report No. 16 Geneva 1985).
4. Visconti, G. *Transport Processes in the Middle Atmosphere* Proc. NATO workshop November 1986, Erice, Sicily (Reidel, Dordrecht, 1987).
5. Butchart, N. & Remsberg, E. E. *J. atmos. Sci.* **43**, 1319–1339 (1986).
6. McIntyre, M. E. & Palmer, T. N. *J. atmos. terr. Phys.* **46**, 825–849 (1984).
7. Charney, J. G. *J. atmos. Sci.* **28**, 1087–1095 (1971).
8. Haynes, P. H. & McIntyre, M. E. *J. atmos. Sci.* (in the press).

9. McIntyre, M. E. & Palmer, T. N. *Pure appl. Geophys.* **123**, 964-975 (1985).
10. Hoskins, B. J., McIntyre, M. E. & Robertson, A. W. Q. *Jl R. met. Soc.* **111**, 877-946 (1985); **113**, 402-404 (1987).
11. Lilly, D. K. *J. atmos. Sci.* **40**, 749-761 (1983).
12. McWilliams, J. C. *J. atmos. Sci.* **42**, 1773-1774 (1985).
13. Clough, S. A., Grahame, N. S. & O'Neill, A. Q. *Jl R. met. Soc.* **111**, 335-358 (1985).
14. Orszag, S. A. *J. atmos. Sci.* **27**, 890-895 (1970).
15. Hoskins, B. J. Q. *Jl R. met. Soc.* **99**, 723-745 (1973).
16. Turner, J. S. *Buoyancy Effects in Fluids*, Ch. 9.1.1 (Cambridge University Press, 1973).
17. Leovy, C. B. *et al. J. atmos. Sci.* **42**, 230-244 (1985).
18. Solomon, S., Garcia, R. R. & Stordal, F. J. *geophys. Res.* **90** 12981-12989 (1985).
19. O'Neill, A. & Pope, V. D. Q. *Jl R. met. Soc.* (submitted).
20. Smith, A. K. *J. atmos. Sci.* **40**, 2484-2496 (1983).
21. Palmer, T. N. & Hsu, C.-P. F. *J. atmos. Sci.* **40**, 909-928 (1983).
22. Saffman, P. G. & Baker, G. R. A. *Rev. Fluid Mech.* **11**, 95-122 (1979).
23. Holloway, G. & Kristmannsson, S. S. *J. Fluid Mech.* **141**, 27-50 (1984).
24. Tung, K. K. In ref. 4: *Transport Processes in the Middle Atmosphere* Proc. NATO workshop, November 1986, Erice, Sicily (Reidel, Dordrecht, 1987).
25. Pitari, G. & Visconti, G. *J. geophys. Res.* **90**, 8019-8032 (1985).
26. Plumb, R. A. & Mahlman, J. D. *J. atmos. Sci.* **44**, 298-327 (1987).
27. Tuck, A. F. *Phil. Trans. R. Soc.* **A290**, 477-494 (1979).
28. Batchelor, G. K. *J. Fluid Mech.* **5**, 113-133 (1959).
29. Cocke, W. J. *Phys. Fluids* **12**, 2488-2492 (1969).
30. Kraichnan, R. H. *J. Fluid Mech.* **64**, 737-762 (1974).
31. Holm, D. D., Marsden, J. E., Ratiu, T. & Weinstein, A. *Phys. Rep.* **123**, 1-116 (1985).
32. McIntyre, M. E. & Shepherd, T. G. *J. Fluid Mech.* **181**, 527-565 (1987).
33. Dritschel, D. G. *J. Fluid Mech.* (submitted).
34. Godson, W. L. *Meteor. Abhandlungen* **36**, 161-206 (1963). See §6.
35. Dobson, G. M. B. *Exploring the Atmosphere* 118 (Clarendon, Oxford, 1963).
36. Farman, J. C. *Phil. Trans. R. Soc.* **B279**, 261-271 (1977).
37. McCormick, M. P., Trepte, C. R. & Kent, G. S. *Geophys. Res. Lett.* **10**, 941-944 (1983).
38. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207-210 (1985).

A small catalytic oligoribonucleotide

Olke C. Uhlenbeck

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA

A 19-nucleotide RNA fragment can cause rapid, highly specific cleavage of a 24-nucleotide RNA fragment under physiological conditions. Because each 19-mer can participate in many cleavage reactions, this molecule has all the properties associated with an RNA enzyme.

CERTAIN phosphodiester bonds in RNA molecules are unusually susceptible to spontaneous cleavage to give 5' hydroxyl and 2',3'-cyclic phosphate termini. In several cases the site of cleavage corresponds to a known site of RNA processing, suggesting that the reaction occurs *in vivo*¹⁻². Perhaps the best characterized examples of such spontaneous cleavage reactions involve certain plant viroids, virusoids and linear satellite RNAs. The mature termini of the 359-nucleotide RNA genome of satellite tobacco ringspot virus (STobRV) can be generated by incubating either naturally occurring dimers or *in vitro* transcripts of cDNA clones in neutral buffers at physiological temperatures^{2,3}. It is believed that this cleavage reaction is the last step in the rolling-circle replication mechanism proposed for STobRV⁴. Similar self-cleavage reactions have been demonstrated for the STobRV (-)strand⁵ and the (+) and (-)strands of avocado sunblotch viroid (ASBV)⁶ and lucerne transient streak virusoid (VLTSV)⁷. Because these self-cleavage reactions are highly specific and show substantial rate enhancements over spontaneous cleavage of RNA, it is clear that specific folding of the RNA molecule is required for the reaction. Recent experiments with deletion mutants have shown that less than 100 nucleotides of the STobRV genome are needed for efficient self-cleavage³. Symons and coworkers⁶⁻⁸ have proposed a 'hammerhead' model for the secondary structure of the self-cleaving domain that is consistent with many of the cleavage sites. This domain consists of three RNA helices and 13 conserved nucleotides that are likely to be involved in further folding to form a tertiary structure (Fig. 1). Although nucleotides on each side of the cleavage site can vary, the position of cleavage with respect to the structure is conserved. Although this model does not provide a rationale for the lability of the bond, it has successfully predicted the site of self-cleavage in a transcript of a new satellite DNA⁹.

Using *in vitro* transcription from synthetic DNA templates, two oligoribonucleotides were synthesized that can combine to form a structure consistent with the consensus self-cleaving domain. Because rapid cleavage of one of the oligomers is observed only when the other is present, the domain is necessary and sufficient for cleavage. The properties of the cleavage reaction have been studied in detail. Nearly complete cleavage occurs even with a large excess of the oligomer that is cleaved. This indicates that the oligomer that is uncleaved can cycle in

the reaction and therefore be considered to act as a catalyst in the cleavage of the other oligomer. At 19 nucleotides, the catalytic oligomer is the smallest ribozyme described to date.

Cleavage with synthetic oligonucleotides

Two oligoribonucleotides were synthesized by transcription of synthetic DNA templates with T7 RNA polymerase (Fig. 1b). The 19-nucleotide fragment (designated 01) and the 24-nucleotide fragment (designated 02) can combine to form the consensus self-cleavage structure without loops I and III (Fig. 1c). The sequence of stem and loop II is identical to that found in the ASBV (-)strand, whereas the sequences of stems I and III were made consistent with an active T7 RNA polymerase promoter¹⁰ and are unlike stems I and III of any known self-cleaving RNA. A major advantage of preparing the self-cleaving RNA structure in two fragments is that, unlike STobRV and VLTSV, cleavage does not occur during the transcription reactions. This permits the isolation and storage of substantial amounts of both RNA fragments.

The cleavage of the combined RNA fragments is shown in Fig. 2a. When equal amounts of 01 and 02 are mixed in 50 mM Tris pH 7.5 and 10 mM MgCl₂ and incubated at 37 °C for one hour, >85% of 02 is cleaved to form a large product P1 and smaller P2. The amount and size of 01 is not altered in the reaction. If 02 is incubated in the absence of 01 or mixed with 01 but not incubated, no cleavage occurs.

Several experiments were performed to demonstrate that products P1 and P2 were the result of cleavage between the C and U residues indicated by the arrow in Fig. 1c and that P1 terminated with a 2',3'-cyclic phosphate. Identification of products was greatly aided by designing 02 such that the cleavage site was the only location where the sequence CpU occurred. Preparations of 02 were synthesized with each of the four α -³²P nucleoside triphosphates and subjected to self-cleavage in the presence of unlabelled 01. The product band P2 is not observed when 02 was made with [α -³²P] UTP. This result is expected, because the labelled phosphate 5' to the only uridine residue in P2 should be attached to P1 as the terminal 2',3'-cyclic phosphate. This assertion is confirmed by the analysis of nuclease P₁ digestion products of each of the labelled P1 molecules. Nuclease P₁ cleaves RNA to give nucleoside 5' monophosphates and does not hydrolyse 2',3'-cyclic phosphates. Only the [α -

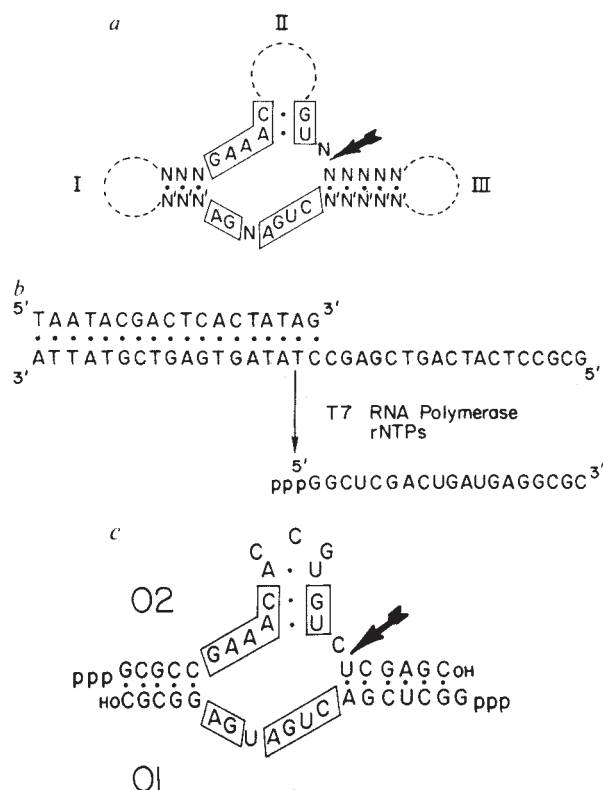


Fig. 1 Secondary structure and synthesis of self-cleaving RNAs. **a**, Consensus self-cleaving domains derived from Hutchins *et al.*⁶ but modified to include information regarding an additional self-cleaving sequence⁹. N, an arbitrary nucleotide complementary to N'. The number of nucleotides in hairpin loops I, II and III vary from 2 to more than 200. Arrow, site of cleavage. **b**, Transcription of a synthetic DNA template by T7 RNA polymerase to produce oligoribonucleotide O1 using the method of Milligan *et al.*¹⁰. **c**, Oligoribonucleotides O1 and O2 annealed to form a secondary structure consistent with the consensus self-cleaving domain. Arrow, putative site of cleavage.

Methods. A typical transcription reaction contained 50 nM DNA template, 3.0 mM each NTP, 18 mM MgCl₂, 1 mM spermidine, 5 mM dithiothreitol (DTT), 40 mM Tris-HCl pH 8.1, 0.01% Triton X-100, 80 mg ml⁻¹ polyethylene glycol 8,000, and 15 units μl⁻¹ T7 RNA polymerase. If ³²P-labelled transcript was needed, 10–100 μCi of the appropriate [α -³²P] NTP was included without changing the total concentration. After incubation for 4 h at 37 °C, reaction products were separated on a 20% polyacrylamide sequencing gel. Approximately equal amounts of O1 and a product one nucleotide longer were observed¹⁰. Oligoribonucleotide O1 was eluted from the gel and precipitated with ethanol. Recovered yield was 100 nmol of O1 per nmol DNA template. Similar yields were obtained with O2.

³²P]UTP- and CTP-labelled P1 molecules show the expected pC>p product (Fig. 2b). Additional evidence for accurate cleavage includes the disappearance of the ribonuclease T₁ digestion product UCUCGp upon self-cleavage and the observation of the predicted nucleotide compositions of P1 and P2 after RNase T₂ digestion (data not shown).

Characterization of the cleavage reaction

The cleavage reaction was studied under a variety of different reaction conditions using labelled O2 and unlabelled O1. Although the rate of reaction varied substantially, the fraction of O2 cleaved at long incubation times remained constant at 85 ± 3%. When the uncleaved O2 was eluted from the gel, mixed with more O1, and incubated again, only about 20% was found to cleave. This experiment suggests that a fraction of the O2 molecules (~12%) are inherently unable to cleave, perhaps as a result of misincorporation during synthesis or transcription

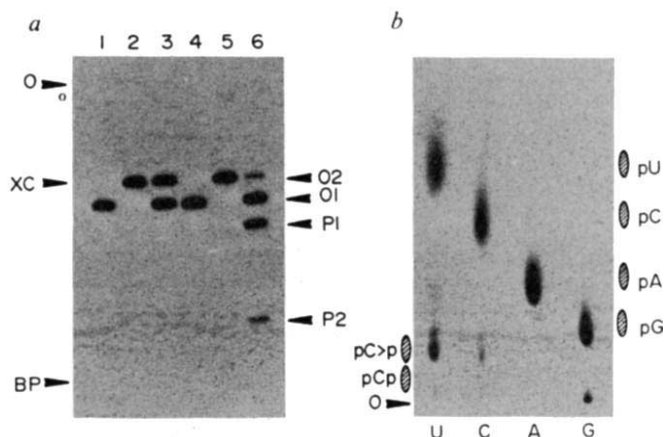


Fig. 2 Cleavage depends upon incubation of both oligonucleotides together and gives a 2', 3'-cyclic phosphate. **a**, Cleavage of the O1-O2 complex. For this, 2.5 pmol of [α -³²P]CTP-labelled O1 (lanes 1 and 4), 2.2 pmol of [α -³²P]CTP-labelled O2 (lanes 2 and 5) or both together (lanes 3 and 6) were incubated in 10 μl of 10 mM MgCl₂ and 50 mM Tris-HCl pH 7.5 for 0 min (lanes 1–3) or 60 min (lanes 4–6) at 37 °C. P1 and P2 are the products of the cleavage reaction (lane 6). Markers: O, top of gel; XC, xylene cyanol; BP, bromophenol. **b**, Nuclease P₁ digestion of product P1 labelled with each of the [α -³²P]-labelled NTPs.

Methods. In **a**, O1 and O2 were heated separately to 90 °C for 1 min in 75 mM Tris-HCl pH 7.5 and allowed to cool to 37 °C. After the addition of MgCl₂ to O2, O1 was added to initiate the reaction. The reactions were stopped by the addition of 10 μl of 10 M urea and 12 mM EDTA. Oligonucleotides were separated on a 20% polyacrylamide/8M urea gel and subjected to autoradiography. Although heat treatment of O1 and O2 was not essential for cleavage, it gave more reproducible results, by minimizing aggregated oligomers that did not enter the gel. In **b**, ~3 pmol of each P1 oligomer was incubated in a 10-μl reaction mix containing 0.3 units P₁ nuclease and 50 mM ammonium acetate pH 7.0 for 30 min at 37 °C. After adding the appropriate 5' nucleotide monophosphate marker, the reactions were spotted on a polyethyleneimine thin-layer plate, developed with 1 M ammonium acetate and subjected to autoradiography. Radioactivity at the origin of the plate was variable and due to incomplete digestion. Positions of non-radioactive markers are indicated on each side of the autoradiogram.

of the synthetic DNA template. Thus, the extent of O1 and O2 cleavage is very high, as in the STobRV(+) and ASBV(-) reactions^{3,6} and in contrast with the ASBV(+) and VLTSV reactions where much less cleavage was observed^{6,7}. To analyse the kinetics of the cleavage reaction (Fig. 3a), the fraction of O2 cleaved at different times was normalized to 0.85 before plotting to give $t_{1/2}$, the time required for 50% cleavage.

The pH dependence of the reaction rate is shown in Table 1. At pH 5 no cleavage is seen, presumably as a result of protonation of adenine and cytosine residues disrupting the tertiary structure of the molecule. By pH 6, cleavage is reasonably rapid and the rate increases slowly with increasing pH. This behaviour differs markedly from the base-catalysed cleavage of RNA, where virtually no cleavage occurs at pH 7 and the rate increases with hydroxide ion concentration as would be expected for a second-order reaction¹¹. The modest dependence of reaction rate on the pH of the buffer for the self-cleavage reaction is quite remarkable. As the hydroxide ion concentration increases 100-fold between pH 7 and pH 9, the reaction rate increases only fourfold. This suggests that the rate limiting step of the reaction does not involve hydroxide ion and that no functional groups involved in the reaction titrate above pH 7.

The cleavage reaction seems to have an absolute requirement for divalent cations (Table 1). When the two oligomers are incubated in 50 mM Tris-HCl, 1 mM EDTA, pH 7.5 no cleavage occurs. Spermidine is not able to replace magnesium ion in the

Table 1 The pH and ionic requirements of cleavage reaction

pH*	$t_{1/2}$ (min)
5.0	—
6.0	20
7.0	8.5
7.5	4.5
8.0	2.6
9.0	2.3
Ions†	
EDTA (1 mM)	—
Spermidine (0.5, 1, 2 mM)	—
NaCl (0.25, 0.5, 1 M)	—
MgCl ₂ (1 mM)	~300
(2.5 mM)	90
(5 mM)	17
(10 mM)	4.4
(20 mM)	3.0
MgCl ₂ (10 mM) + 1 M NaCl	4.5
MnCl ₂ (10 mM)	~0.6
CaCl ₂ (10 mM)	15.0
ZnCl ₂ (10 mM)	—

* Buffer: 50 mM Tris-HCl (pH 7-9) or 50 mM K⁺ MES (pH 5-6) and 10 mM MgCl₂ at 37 °C.

† Buffer: 50 mM Tris-HCl pH 7.5 at 37 °C. —, No cleavage observed after 6 h.

reaction as it can in the self-cleavage of STobRV². The reaction rate was found to increase steadily with increasing MgCl₂ concentration until a maximum is reached at ~20 mM MgCl₂. Because MgCl₂ can stabilise RNA helices in this same concentration range, it is possible that the divalent ion simply stabilizes the complex between the two oligomers. However no cleavage is observed in 1 M NaCl which is generally equally effective in stabilizing RNA secondary structure. Because the same MgCl₂ concentrations are required for rapid cleavage in the presence of 1 M NaCl, the experiments suggest that magnesium ion has a special function in the reaction. Both Mn²⁺ and Ca²⁺, but not Zn²⁺ can substitute for Mg²⁺ in the reaction.

The rate of cleavage was measured at a number of temperatures. An Arrhenius plot (Fig. 3b) is linear between 10 °C and 37 °C and indicates an activation energy of 13.1 cal mol⁻¹. Above 37 °C, the rate of the reaction increases more slowly until the reaction reaches a maximal rate at 50 °C and then decreases at higher temperatures. These data suggest that the RNA structure essential for cleavage is unable to form at the higher temperatures. This behaviour is very similar to protein catalysed enzyme reactions although the temperature range over which denaturation occurs is broader than that generally encountered.

Catalytic cleavage

Because O1 is not consumed in the cleavage reaction, it should be able to promote cleavage of additional O2 molecules if the products can dissociate from O1 and a new O1-O2 complex can form. An experiment demonstrating multiple cleavage of O2 molecules by O1 is shown in Fig. 4. A 21-fold molar excess of O2 was incubated with O1 at 55 °C, a temperature where the rate of cleavage remains high but product dissociation would be favoured. Because >90% of O2 is cleaved in three hours, each O1 molecule has participated in an average of 20 cleavage events. The rate of cleavage is initially about 4 pmol min⁻¹ which is comparable to the rate at 55 °C seen in Fig. 3 when equal concentrations of O1 and O2 were used. Thus, there is no indication that the rate of cleavage changes after a single turnover occurs, suggesting that product dissociation is not rate limiting for cleavage. Either P1 and P2 directly dissociate from O1 after cleavage or a new O2 molecule displaces P1 and P2 in a strand exchange reaction.

At longer incubation times, a slower rate of cleavage is observed so that about two hours is required to cleave the last

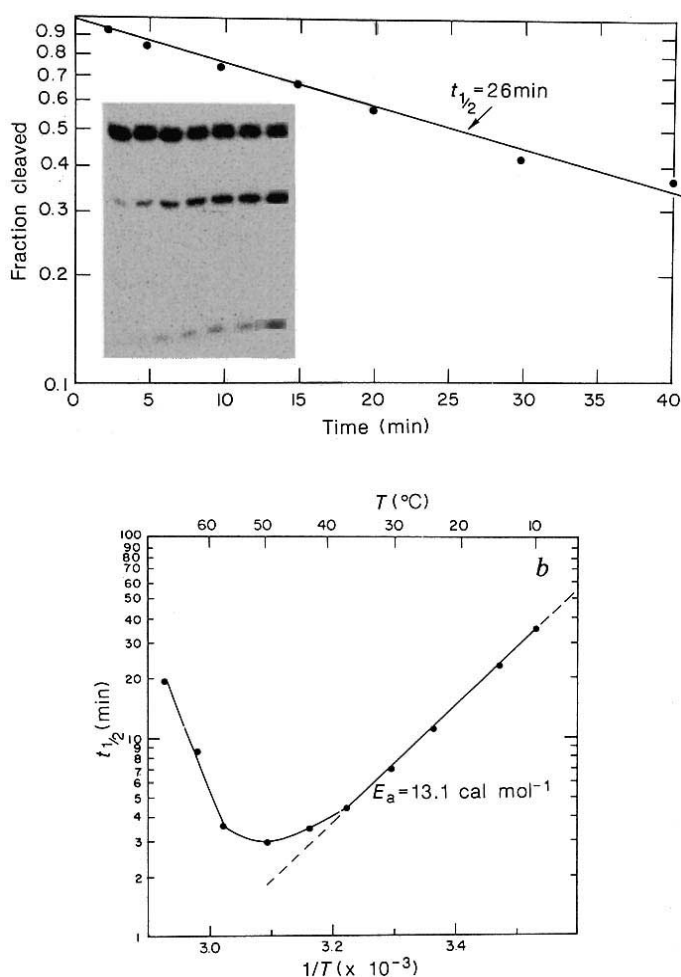


Fig. 3 Temperature dependence of cleavage rate. *a*, Cleavage kinetics at 16 °C: 10 pmol [α -³²P] CTP-labelled O2 (2 nCi pmol⁻¹) was incubated with 15 pmol unlabelled O1 at 16 °C in 40 μ l of 10 mM MgCl₂ and 50 mM Tris-pH 7.5 and 5- μ l aliquots removed at the indicated times. Products of the cleavage reactions were separated on a 20% polyacrylamide/8 M urea gel, localized by autoradiography (inset) and counted. The fraction of O2 cleaved was normalized to 85% total cleavage before plotting. *b* Arrhenius plot of the cleavage reaction. The dashed line has a slope (activation energy) of 13.1 kcal mol⁻¹.

50 pmol (Fig. 4). This suggests that the reaction rate is influenced by the concentration of reactants and products. To test this, the initial rate of cleavage at varying concentrations of O2 was determined at a constant O1 concentration (0.1 μ M). The initial rate of cleavage increases with O2 concentration until an upper limit is reached. This suggests that a Michaelis-Menten type mechanism can be used to evaluate the data. An Eadie-Hofstee plot of the initial rate data (Fig. 5a) gives a reasonable fit with $K_m = 0.62 \mu$ M and $k_{cat} = 0.5 \text{ min}^{-1}$. At a high, near-saturating concentration of O2 (10.4 μ M), the initial rate of cleavage is directly proportional to the concentration of O1 added over a 25-fold range (Fig. 5b), as would be expected for an enzyme at saturating substrate concentration.

Discussion

The sequences of the two RNA fragments studied in this work are unlike those in any known natural RNA, but conform to the model proposed by Symons and coworkers⁶⁻⁸ for a self-cleaving domain. The observation of rapid cleavage provides additional support that the combination of nucleotide sequence and secondary structure shown in Fig. 1a is sufficient for

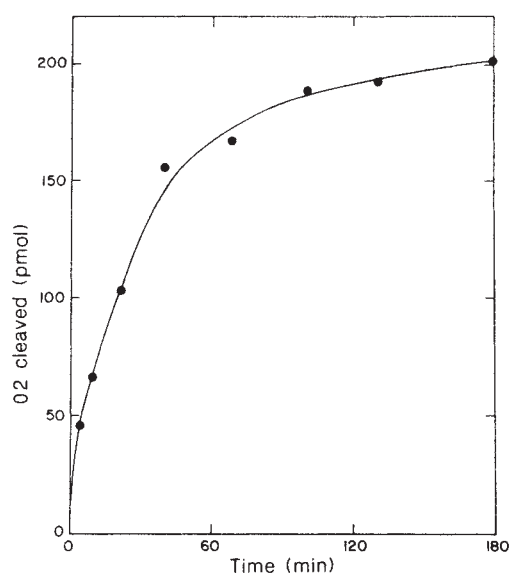


Fig. 4 Multiple cleavage of O2 molecules by O1: 208 pmol of ^{32}P -labelled oligoribonucleotide O2 were mixed with 10 pmol of O1 in 40 μl of 50 mM Tris-*pH* 7.5 and 10 mM MgCl_2 and incubated at 55 $^{\circ}\text{C}$ for the indicated times. The products were separated and analysed as in Fig. 3. No cleavage was observed without oligonucleotide O1.

cleavage. Considering that the RNA chain must fold into a conformation that destabilizes a certain phosphodiester bond, the size of the self-cleaving domain is remarkably small. Molecules O1 and O2 combined contain only 43 nucleotides and it is likely that two base pairs could be removed from both helix I and helix III and one base pair from helix II, making a minimal active structure of 33 nucleotides or less.

The Symons model should allow the prediction of other sites of self-cleavage based solely on nucleic acid sequence. However because only a limited number of self-cleaving domains are currently known, it is unclear how many features of the model are essential for the cleavage reaction. For example, not all of the conserved nucleotides may be needed for cleavage, or it may be possible to insert additional nucleotides into the domain without altering cleavage. By preparing sequence variants of one or both oligomers it will be possible to explore the sequence and structure requirements for the cleavage reaction.

Unlike all known natural examples of self-cleaving RNAs, the self-cleaving domain described here is constructed from two separate RNA fragments. Thus, what is normally an intramolecular reaction has been converted to an intermolecular reaction. This resembles the work of Zaug and Cech¹² who showed that the intramolecular transesterification reaction catalysed by the *Tetrahymena* intervening sequence can be converted to an intermolecular reaction by separating the substrate from the catalytic domain of the intron. Although the motivation for preparing the self-cleaving domain as two separate fragments was primarily to simplify the study of the reaction, the demonstration of active cleavage in *trans* may have biological significance as well. For example, an RNA molecule containing the O1 sequence could act as a nuclease by reversibly combining with O2 sequences embedded in a number of different RNAs. A self-cleaving mechanism involving two RNA fragments may also explain why RNAs presumed to undergo processing from multimers to monomers *in vivo* contain only part of the consensus self-cleaving domain and cleave very poorly *in vitro*¹³⁻¹⁵. A cellular RNA component that could act catalytically may be needed to complete the self-cleaving structure. The ability of small RNAs to act as sequence specific nucleases would be an interesting addition to the available RNA processing mechanisms. In many respects this sort of mechanism resembles nuclear

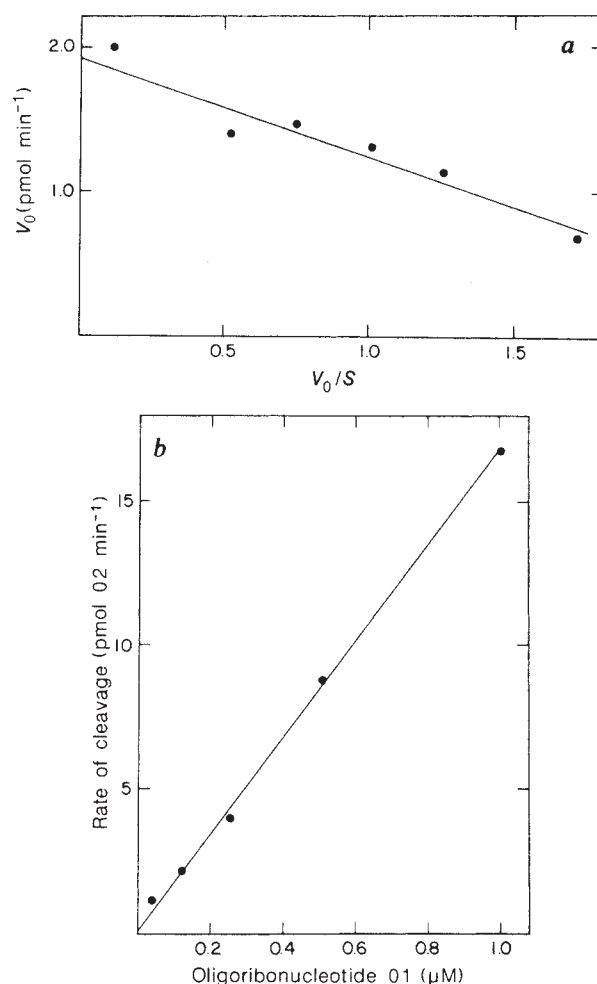


Fig. 5. Kinetics of catalytic cleavage. *a*, Eadie-Hofstee plot of kinetic data in 50 mM Tris-HCl *pH* 7.5, 10 mM MgCl_2 at 55 $^{\circ}\text{C}$. Initial rates of cleavage were determined as in Fig. 4 with 4 pmol O1 and 16-640 pmol of O2. The line fits the data with $K_m = 0.63 \mu\text{M}$ and $V_{\text{max}} = 1.9 \text{ pmol min}^{-1}$. *b*, Initial rate of cleavage of 10.4 μM oligoribonucleotide O2 with 0.04-1 μM O1 in 50 mM Tris-HCl *pH* 7.5, 10 mM MgCl_2 at 55 $^{\circ}\text{C}$.

messenger RNA splicing where sequences in small nuclear RNAs combine with pre-mRNA sequences to form the splicing substrate¹⁶.

At *pH* 7.5 and 37 $^{\circ}\text{C}$, the half-life of an equimolar mixture of O1 and O2 was found to be about 5 min. The rate of random hydrolysis of RNA under these conditions is much slower and difficult to determine accurately. Early estimates of $t_{1/2}$ at 2-15 hours for cleavage of tobacco mosaic virus RNA¹⁷⁻¹⁸ may be inaccurate owing to trace amounts of contaminating ribonuclease. Incubation of O2 by itself for 24 h led to no detectable hydrolysis. A reasonable estimate is that at least a 1,000-fold rate enhancement of phosphodiester bond cleavage is achieved by the self-cleaving domain.

Considering that the rate of strand exchange of short RNA helices near T_m , the melting temperature, is quite rapid, it is not surprising that a single O1 molecule can participate in the cleavage of many O2 molecules. A mechanism can be proposed where chain cleavage is followed by a new O2 displacing the products. The slower of these two steps will limit the overall cleavage rate and result in the observed saturation of the rate of cleavage at high initial O2 concentrations. Although further kinetic experiments will be required to confirm a mechanism for the reaction, it is clear that the cleavage of O2 by O1 resembles an enzyme reaction. It is interesting to note that the k_{cat} of 0.5 min^{-1} for the O1-O2 cleavage reaction is very similar to the

values reported for other ribozyme reactions. The *Tetrahymena* intervening sequence has a $k_{\text{cat}} = 2 \text{ min}^{-1}$ with a short oligomer substrate¹² and *Escherichia coli* M1 RNA cleaves a tRNA precursor with an optimal k_{cat} of 1 min^{-1} (ref. 19). These values are low compared to most protein enzymes, but are similar to the value for some, such as the restriction endonuclease *EcoRI* (ref. 20).

Enzymes are generally thought to show a high degree of substrate specificity, to enhance substantially the reaction rate and to act many times without being consumed in the reaction. It is clear that the 19-nucleotide RNA fragment 01 has all these properties when it participates in the cleavage of the 24-nucleotide 02 fragment. The reaction is highly specific, because 01 catalyses the cleavage of only one of the phosphodiester bonds in 02. We have shown that 01 is not capable of cleaving a variety

of other RNAs and therefore is not a general nuclease (data not shown). As described above, 01 acts under physiological conditions to enhance the rate of 02 cleavage by at least three orders of magnitude. Finally, 01 is capable of cycling many times without being consumed in the reaction. The action of 01 on its substrate 02 does not resemble most protein enzyme reactions in the sense that large rearrangements in the RNA structure are likely to occur during the reaction. The helices joining the two molecules must break and reform during each turnover. The formation and breakage of a short helix between ribozyme and substrate is also a characteristic of reactions involving the *Tetrahymena* intron²¹.

I thank J. Milligan and D. Groebe for teaching me oligoribonucleotide synthesis and T. Cech for useful discussions. This work was supported by grants from NIH.

Received 1 June; accepted 29 June 1987.

1. Watson, N., Gurevitz, M., Ford, J. & Apirion, D. *J. molec. Biol.* **172**, 301–323 (1984).
2. Prody, G. A., Bakos, J. T., Buzayan, J. M., Schneider, I. R. & Bruening, G. *Science* **231**, 1577–1580 (1986).
3. Buzayan, J. M., Gerlach, W. L. & Bruening, G. *Proc. natn. Acad. Sci. USA* **83**, 8859–8862 (1986).
4. Branch, A. D., Robertson, H. D. & Dickson, E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6381–6385 (1981).
5. Buzayan, J. M., Gerlach, W. L. & Bruening, G. *Nature* **323**, 349–352 (1986).
6. Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. *Nucleic Acids Res.* **14**, 3627–3640 (1986).
7. Forster, A. C. & Symons, R. H. *Cell* **49**, 211–220 (1987).
8. Keese, P. & Symons, R. H. in *Viroids and Viroid-Like Pathogens* (ed. Semancik, J. S.) 1–47 (CRC Press, Boca Raton, Florida, 1987).
9. Epstein, L. M. & Gall, J. G. *Cell* **48**, 535–543 (1987).
10. Milligan, J., Groebe, D., Witherell, G. & Uhlenbeck, O. *Nucleic Acids Res.* (submitted).
11. Bock, R. M. *Meth. Enzym.* **12A**, 218–221 (1967).
12. Zaug, A., & Cech, T. R. *Science* **231**, 470–475 (1986).
13. Robertson, H. D., Rosen, D. L. & Branch, A. D. *Virology* **142**, 442–447 (1985).
14. Visvader, J. E., Forster, A. C. & Symons, R. H. *Nucleic Acids Res.* **13**, 5843–5856 (1985).
15. Wang, K. et al. *Nature* **323**, 508–514 (1986).
16. Maniatis, T. & Reed, R. *Nature* **325**, 673–678 (1987).
17. Ginoza, W. *Nature* **181**, 958–961 (1958).
18. Eigner, J., Boedtker, H. & Michaels, G. *Biochem. biophys. Acta* **51**, 165–168 (1961).
19. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849–857 (1983).
20. Modrich, P. & Zabel, D. *J. biol. Chem.* **251**, 5866–5871 (1976).
21. Cech, T. R. & Bass, B. A. *Rev. Biochem.* **55**, 599–629 (1986).

LETTERS TO NATURE

Heliomagnetic dipole moment and daily variation of cosmic rays underground

K. Nagashima, H. Ueno & K. Fujimoto

Faculty of Science, Nagoya University, Nagoya 464, Japan

Since the discovery of Forbush decrease of cosmic-ray intensity, it has become clear that the intensity and its daily variation are subject to various kinds of solar modulation; these are due to the diffusion-convection^{1,2} of cosmic rays in interplanetary magnetic fields, which is controlled by various kinds of solar activity. We now demonstrate that the amplitude of diurnal variation of cosmic-ray intensity is intimately correlated with the magnitude of magnetic dipole moment of the Sun and that, contrary to the conventional concept of the modulation, such a correlation exists only in the cosmic-ray high-rigidity region.

Cosmic-ray data used for this analysis are those observed by muon telescopes at Sakashita underground station^{3,4} from 1978 to 1985. Some of characteristics of the telescopes are listed in Table 1. Amplitudes (A) of the monthly averaged diurnal variations of cosmic rays have been obtained after applying the barometer correction to these data. The heliomagnetic dipole moment (M) to be compared with the amplitude is defined by

$$|M| = R_{\odot}^3 \sqrt{(g_{1,0}^2 + g_{1,1}^2 + h_{1,1}^2)}$$

where $R_{\odot}$ is the solar radius, and $g_{1,0}$, $g_{1,1}$ and $h_{1,1}$ are coefficients of the first-order surface harmonic of the magnetic scalar poten-

tial (Ψ) in the heliographic polar coordinate (r, θ, ϕ)⁵:

$$\begin{aligned} \Psi(r, \theta, \phi) = & \sum_{n=1}^{\infty} \sum_{m=0}^n \alpha_n(r) \\ & \times (g_{n,m} \cos m\phi + h_{n,m} \sin m\phi) P_{n,m}(\cos \theta) \\ \alpha_n(r) = & (R_{\odot}^{n+2}/r^{n+1}) \{1 - (r/R_{\odot})^{2n+1}\} \quad \text{for } R_{\odot} \geq r \geq R_{\odot} \end{aligned}$$

where $P_{n,m}(\cos \theta)$ is the seminormalized spherical function⁶. $R_{\odot}$ is the radius of outer boundary called the source sphere⁵ and assumed to be $2.5R_{\odot}$. This potential is arranged so that θ and ϕ components (B_{θ} , B_{ϕ}) of the magnetic field ($\mathbf{B} = \mathbf{B}_r + \mathbf{B}_{\theta} + \mathbf{B}_{\phi}$) tend to zero at $r = R_{\odot}$ and, therefore, does not exactly express the usual dipole field by the terms with $n=1$. It is noted, however, that at the vicinity of $R_{\odot}$, the deviation from the dipole field is only $\sim (R_{\odot}/R_s)^3$ and one can approximately define the magnetic dipole moment of the Sun. This moment is equivalent to the r.m.s. of the field $\sqrt{\langle B_r^2 \rangle}$ for $n=1$ on the source surface, which was used by Hoeksema and Scherrer⁵ for representing the time-variation of the field as $M = R_s^3 \sqrt{\langle B_r^2 \rangle}/3$. The coefficients $g_{n,m}$ and $h_{n,m}$ for $m \leq n \leq 9$ have been presented by Hoeksema and Scherrer⁵ every half period of Carrington rotation from 1976 to 1985. We use these coefficients for $n=1$ after averaging those in the first and second half periods of each rotation. Figure 1 shows the magnetic moment M together with the amplitude $A_v(\text{Sak})$ for the vertical telescope. For reference, the sunspot number (R_z) and also the inclination $|\lambda_M|$ of the dipole axis with respect to the solar equatorial plane ($\theta = 90^\circ$), defined by

$$|\lambda_M| = \arctan [|g_{1,0}| / \sqrt{(g_{1,1}^2 + h_{1,1}^2)}]$$

are shown. This quantity $|\lambda_M|$ is supposed to be related to the inclination Λ of the heliomagnetospheric neutral sheet in space relative to the solar equatorial plane, as $\Lambda = 90^\circ - |\lambda_M|$.

It is clearly seen in Fig. 1 that the time variation of $A_v(\text{Sak})$ has an intimate correlation with M . Such a correlation disappears for the amplitudes in lower rigidity region such as that ($A_v(\text{Nag})$) observed by a vertical muon telescope (median rigidity $P_m \sim 60\text{GV}$) at Nagoya surface station⁷ (35°N , 137°E) (see

Table 1 Muon telescopes at Sakashita underground station

Location	Depth (hg cm ⁻²)	Telescope		Flux ($\times 10^3$ h)	Median rigidity P_m (GV)
		Zenith/Azimuth			
36° N	80	V	0/—	371	331
138° E	80	S2	60°/166°	27	540

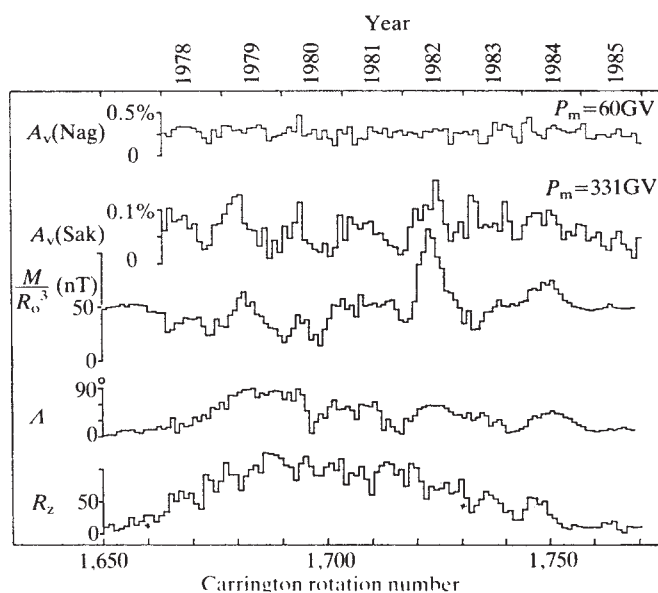


Fig. 1 Time variations of amplitudes A_v (NAG) and A_v (Sak) of diurnal variation of intensity of cosmic rays observed by vertical muon telescopes at Nagoya surface station and at Sakashita underground station, the heliomagnetic dipole moment (M), the inclination (A) of the heliomagnetospheric neutral sheet with respect to solar equatorial plane and the sunspot number (R_z). Upper and lower abscissas express, respectively, the year and the Carrington rotation number; P_m is the median rigidity of cosmic rays. Errors of the amplitudes estimated from counting rates are $3 \times 10^{-3}\%$ for A_v (Nag) and $9 \times 10^{-3}\%$ for A_v (Sak).

Fig. 1). The amplitude of neutron intensity ($P_m \sim 30$ GV) also does not have such a correlation with M . On the other hand, the uppermost rigidity for the correlation is not certain but is supposed to be ≤ 540 GV as the amplitude A_{S2} (Sak) for S2-telescope in Table 1 does not have any correlation with M . We call this phenomenon the preferential correlation of the amplitude with the magnetic dipole moment of the Sun.

Generally, almost all the solar modulations are supposed to be more dominant in lower rigidity region but this phenomenon is quite contrary. In this respect, the preferential correlation can be regarded as an anomalous phenomenon, which cannot be explained by the hypothesis^{8,9,10} that the intensity and/or its diurnal variation is intimately related to the inclination A of the neutral sheet because, as can be seen in the figure, the amplitude A_v (Sak) does not correlate with A throughout all of the period in question. Although it cannot be explained by any other theories or hypotheses so far presented, it seems to provide an important clue for understanding the interaction of cosmic rays with the heliomagnetosphere from the following points of view: (1) the amplitude directly and almost quantitatively correlates with the physical quantity (magnetic moment) not activity index of the Sun; (2) the correlation is dominant only in a high-rigidity region. These two characteristics seem to suggest that cosmic rays interact more directly with the magnetic dipole field of the Sun with the increase of their collision mean free path in the heliosphere.

Finally, it is noted that cosmic ray semidiurnal variation also seems to have such a preferential correlation with M . Details will be reported elsewhere.

We are grateful to Mr T. Yamada for the maintenance of Sakashita underground station, and to Mr M. Orito, Mrs H. Satake and Miss S. Demura for their help in processing data and preparing the manuscript.

Received 2 March; accepted 12 May 1987.

1. Parker, E. N. *Interplanetary Dynamical Processes* 161-243 (Interscience, New York, 1965).
2. Gleeson, L. J. & Axford, W. I. *Astrophys. J.* **149**, L115-L118 (1967).
3. Ueno, H., Yamada, T., Fujii, Z., Fujimoto, K., Kondo, I. & Nagashima, K. *Proc. Int. Conf. Cosmic Rays*, Kyoto, Vol. 4, 180-185 (1979).

4. Ueno, H., Yamada, T., Fujii, Z., Sakakibara, S., Fujimoto, K. & Nagashima, K. *Report of Cosmic-Ray Research Laboratory*, no. 5 ix-xx (Univ. Nagoya, 1981).
5. Hoeksema, J. T. & Scherrer, P. H. *Report UAG-94, World Data Center A for Solar-Terrestrial Physics* (January 1986).
6. Chapman, S. & Bartels, J. *Geomagnetism*, Vol. 2, 617-624 (Oxford University Press, London, 1940).
7. Nagashima, K., Fujii, Z., Sakakibara, S., Fujimoto, K. & Ueno, H. *Report of Cosmic-Ray Research Laboratory*, no. 10, ix-xx (Univ. Nagoya, 1984).
8. Swinson, D. B., Saito, T. & Mori, S. *J. geophys. Res.* **86**, 8845-8851 (1981).
9. Saito, T. & Swinson, D. B. *J. geophys. Res.* **91**, 4536-4542 (1986).
10. Swinson, D. B. & Saito, T. *J. geophys. Res.* **91**, 13675-13678 (1986).

Superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystals

L. F. Schneemeyer, J. V. Waszczak, T. Siegrist, R. B. van Dover, L. W. Rupp, B. Batlogg, R. J. Cava & D. W. Murphy

AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

Increased understanding of fundamental properties of the new cuprate superconductors depends on advances in the preparation of quality materials for study. Early work focused on the study of sintered polycrystalline samples. In some cases, grain growth during sintering yielded individual grains large enough ($\sim 80 \mu\text{m}$) for single crystal X-ray determination¹. Larger single crystals allow measurements of physical properties including measurements of anisotropic behaviour. The anisotropy of upper critical fields in $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystals has been reported by Iye *et al.*² who found critical fields characteristic of a quasi-two-dimensional superconductor, which are highest when the field is orientated perpendicular to the c -axis. Dinger *et al.*³ have measured critical fields and critical currents of single crystals and find anisotropies of 10 and greater³. These studies contain little detail on the crystal growth of $\text{YBa}_2\text{Cu}_3\text{O}_7$. Here we report the growth of single crystals as large as 4 mm of $\text{YBa}_2\text{Cu}_3\text{O}_7$ and $\text{DyBa}_2\text{Cu}_3\text{O}_7$ and confirm superconductivity in them by magnetic and electrical measurements. Single crystals of other rare earth barium cuprates are also obtained although we have yet to optimize growth conditions for these phases.

The high- T_c material, $\text{YBa}_2\text{Cu}_3\text{O}_7$, does not melt congruently, precluding many melt growth techniques for crystal growth. Also, limited solubility in, or reactions with, common flux materials limits standard flux growth techniques as well. Our

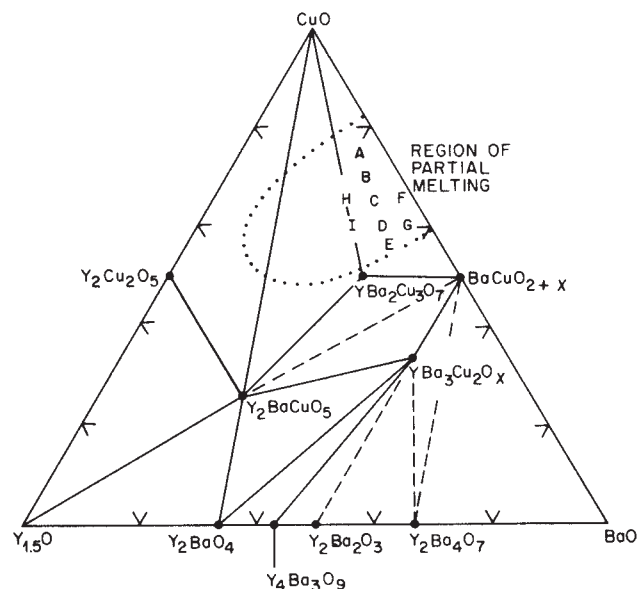


Fig. 1 $\text{BaO-CuO-Y}_2\text{O}_3$ ternary phase diagram⁴ (~ 950 - $1,000^\circ\text{C}$). Dotted line represents region of partial melting.

approach to crystal growth takes advantage of the unusual phase diagram⁴ near 1,000 °C shown in Fig. 1. The method relies on the existence of a region of partial melting near the superconducting line phase, $\text{YBa}_2\text{Cu}_3\text{O}_7$. Mixtures in this region, rich in Cu and Ba, serve as a flux for crystal growth. A similar process was used to obtain strontium lanthanum copper oxide crystals with the K_2NiF_4 structure from Sr-La-Cu-O melts⁵. In a typical experiment to grow $\text{YBa}_2\text{Cu}_3\text{O}_7$ crystals, a mixture corresponding to one of the compositions labelled A-I in the region of partial melting is used. Experimental conditions are summarized in Table 1. The mixture is heated to 800 °C over ~4 h, then to

Table 1 Conditions for the growth of $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystals

Sample no.	M:Cu	Ba:Y	Melt composition	Comments
A	1:3	4:1	$\text{YBa}_4\text{Cu}_{15}\text{O}_x$	Cu-poor
B	5:12	4:1	$\text{YBa}_4\text{Cu}_{12}\text{O}_x$	Good xls
C	1:2	4:1	$\text{YBa}_4\text{Cu}_{10}\text{O}_x$	Best xls
D	2:3	4:1	$\text{Y}_2\text{Ba}_8\text{Cu}_{15}\text{O}_x$	Good morphology
E	3:4	4:1	$\text{Y}_3\text{Ba}_{12}\text{Cu}_{20}\text{O}_x$	Few xls
F	1:2	9:1	$\text{YBa}_9\text{Cu}_{20}\text{O}_x$	CuO xls
G	2:3	9:1	$\text{Y}_2\text{Ba}_{18}\text{Cu}_{30}\text{O}_x$	CuO xls
H	1:2	2:1	$\text{YBa}_2\text{Cu}_6\text{O}_x$	CuO xls
I	2:3	2:1	$\text{Y}_2\text{Ba}_4\text{Cu}_9\text{O}_x$	Small xls

1,000 °C at 50° per hour, held at 1,000 °C for 3–5 h, and slow-cooled to a temperature below the melting temperature (~880 °C) at rates varying from 4 to 15° per hour. In general, larger crystals were obtained with slower cooling rates. Because we found no solvent for flux removal that did not also destroy the crystals, the crystals were mechanically separated from the flux. Crystals often protruded from the hardened mixture, appearing to grow out of the melt, facilitating their isolation. In CuO-rich melts, single crystals of CuO having a different growth habit than the perovskite crystals were sometimes isolated. These pyramidal CuO crystals could be large, up to several millimetres in length.

A variety of factors were explored to optimize crystal growth. In addition to melt composition and cooling rate, several types of crucibles were used including Ni, Pt, alumina, magnesia and zirconia. Some degree of contamination of the crystals resulted from most crucible materials. Koors high-alumina-wave crucibles were used for the majority of the experiments and yielded crystals with $T_c \sim 90$ K. However, many crystals are doped with aluminium, leading to lowered transition temperatures. The best results were obtained with ThO_2 crucibles. Apparently the Th^{4+} ion is too large to enter the $\text{YBa}_2\text{Cu}_3\text{O}_7$ lattice.

The crystals form with a layered-type morphology, similar to mica, but generally with squared corners. The *c*-axis is normal to the platelets and the aspect ratio is ~10. The crystals with highest transition temperatures have well-defined crystal faces. A $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystal measuring 3.7 mm is shown in Fig. 2. Morphology and crystal size were strongly dependent on melt composition. Crystals from more copper-rich melts were thinner with larger surface areas. Crystals from very copper-rich melts (Fig. 1 *a*) had lowered superconducting transition temperatures and were heavily contaminated by impurities from all crucibles used.

The crystals are bulk superconductors as demonstrated by d.c. diamagnetic susceptibility. As is the case for ceramic samples^{6,7}, transition temperatures are strongly dependent on the oxygen stoichiometry. Figure 3 shows the temperature dependence of the susceptibility for a single $\text{YBa}_2\text{Cu}_3\text{O}_x$ crystal as grown in air (oxygen-deficient) and following an anneal in oxygen at 500 °C for several hours. The oxygen anneal raises the superconducting onset temperature and sharpens the transition. The magnitude of the Meissner signal is unchanged by the oxygen anneal which implies that in both cases the sample is uniformly superconducting. We point out the dip in the susceptibility signal at ~60 K in the annealed sample. We speculate that

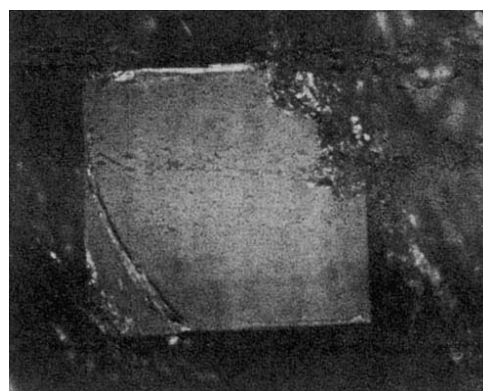


Fig. 2 Photograph of a $\text{YBa}_2\text{Cu}_3\text{O}_7$ crystal measuring 3.7 mm on an edge.

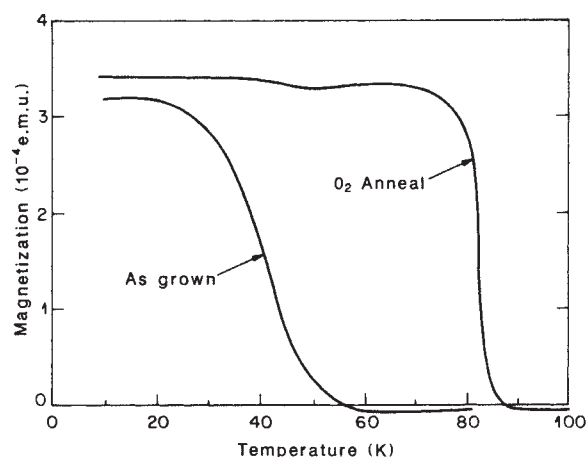


Fig. 3 The temperature dependence of the diamagnetic magnetization of a $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystal as grown and following an oxygen anneal as described in the text. The low-temperature magnetization represents an exclusion of ~90% of the magnetic flux from the crystal. Data include an instrumented offset of -1×10^{-6} e.m.u.

this may correspond to a very small portion of the sample which is not fully oxygenated and therefore shows the lower transition temperature.

Superconductivity was also confirmed by resistivity measurements. Figure 4 shows the resistivity as a function of temperature for an oxygen-annealed single crystal of ~2 mm² surface area and 0.5 mm thickness with four indium contacts soldered in the van der Pauw configuration on an *a*-*b* face. The resistive onset occurs at 91 K with $R = 0$ at 88 K. The contact resistance of the (~0.1 mm²) contacts was ~10³ Ω, substantially higher than that obtained for soldered indium contacts on sintered ceramic $\text{YBa}_2\text{Cu}_3\text{O}_7$ (<1 Ω mm⁻²), although in the latter case the actual contact area is ill-defined. The normal-state resistivity obtained in this measurement is ~5,000 μΩcm, which is a complicated average of the presumably anisotropic resistivities. It is not

Table 2 Crystallographic data for $\text{M}_3\text{Cu}_3\text{O}_x$

M_3	Comments	<i>a</i> Å	<i>b</i> Å	<i>c</i> Å
YBa_2	as grown	3.8605(5)	3.8613(5)	11.718(1)
YBa_2	annealed	3.854(1)	3.861(1)	11.675(2)
DyBa_2	as grown-xl no. 1	3.857(1)	3.859(1)	11.708(5)
DyBa_2	as grown-xl no. 2	3.856(1)	3.858(1)	11.704(4)
DyBa_2	annealed-xl no. 2	3.846(2)	3.856(2)	11.664(7)

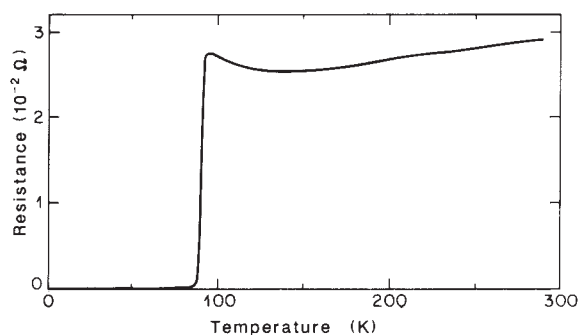


Fig. 4 Temperature dependence of the resistivity of an oxygen-annealed $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystal measured along the b -axis. Sample dimensions are $1.2 \times 0.8 \times 0.5 \text{ mm}^3$.

certain why the resistivity of this sample depends so weakly on temperature. Measurements of additional crystals are under way to determine whether this behaviour is intrinsic, due to slight deviations in metal stoichiometries, or extrinsic, due perhaps to incomplete annealing.

The crystals as grown are slightly oxygen-deficient. Lattice constants obtained from the same crystals before and after oxygen annealing are given in Table 2 for the yttrium and dysprosium phases. Absolute 2θ -values in the range of 50° – 65° of at least 20 reflections were obtained on a CAD-4 κ -axis diffractometer. Within the limits of the errors, the crystals as grown show tetragonal metric symmetry with an elongated c -axis. Upon annealing, the c -axis contracts and, despite twinning and an enlarged mosaic, the orthorhombic distortion becomes clearly observable. Due to microtwinning the difference in the a and b lattice parameters is not as pronounced in these crystals as has previously been observed in ceramic powders by powder X-ray diffraction⁶.

In the case of the Y-Ba-Cu-O mixtures examined here, complete melting does not occur. Thus, this is not a true flux growth. Rather, partial melting occurs and crystal growth is probably a complicated process which depends not only on melt composition and cooling rate, but also on factors such as the rate of heating as well. Indeed, high-quality crystals are obtained even when very little melting has occurred.

The existence of low-melting Y-Ba-Cu-O mixtures in the vicinity of the $\text{YBa}_2\text{Cu}_3\text{O}_7$ phase may account for the growth of large grains under some sintering conditions. The $\text{YBa}_2\text{Cu}_3\text{O}_7$ phase may partly decompose at high sintering temperatures yielding material which acts as a flux to promote grain growth. While such a mechanism is favourable for the growth of crystals of the superconducting phase, it may suggest complications in ceramic processing leading to poorly controlled grain-size distribution and inactive material at the grain boundaries.

The availability of single crystals is critical to many fundamental studies of high-temperature superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_7$ -type phases. Data correlating structure-property relationships may hold the key to understanding the nature of superconductivity in this fascinating class of materials.

We thank S. A. Sunshine, J. P. Remeika, P. M. Bridenbaugh, R. A. Laudise and R. M. Fleming for helpful discussions.

Improved low contact resistance in high- T_c Y-Ba-Cu-O ceramic superconductors

J. van der Maas*, V. A. Gasparov*[‡] & D. Pavuna†

* Institut de Physique Expérimentale de l'Université de Lausanne, CH-1015 Lausanne, Switzerland

† Institute of Applied Physics, Federal Institute of Technology, CH-1015 Lausanne, Switzerland

One of the important problems for electrical applications of high-temperature superconductors^{1–3} is the contact resistance between the superconductor and a normal metallic wire. Recently, AT&T Bell researchers⁴ reported that the contact technique limited measurements of a critical current in their Y-Ba-Cu-O samples: even after necking the cross-section down to $2 \times 10^{-3} \text{ cm}^2$, the sample stayed superconducting, the current through the contacts being limited to $\sim 2 \text{ A}$. Apparently contacts are a real problem, because the very high critical-current densities as reported recently by IBM laboratories⁵ have been measured on films with contact resistances of $\sim 20 \Omega$, and a pulsed-current method had to be used. Here we report a simple method of contact preparation for high- T_c copper-oxide materials (Y-Ba-Cu-O). Based on high-temperature metallization it reduces the contact resistance from typically $1 \Omega \text{ mm}^{-2}$ down to $\leq 10 \mu\Omega \text{ mm}^{-2}$ (77 K), and enables us to measure the critical current of a bulk sample, with current densities in the contact area in excess of $2,000 \text{ A cm}^{-2}$.

Our bulk material is prepared in $\text{YBa}_2\text{Cu}_3\text{O}_7$ nominal composition by a standard powder method⁴. After annealing in air, sample bars and plates are cut out of the prepared pellet and silver epoxy contacts are painted on the Y-Ba-Cu-O samples before the final annealing at 900°C in streaming oxygen. The type of silver epoxy used (epo-tek H20S) seems not to be critical because at these high temperatures the solvent of the glue completely disappears and one is left with a pure silver film contact. Silver wires (0.12 mm diameter) are attached to the contact surface with silver epoxy glue; the principal advantage of pure silver leads being that the sample can be repeatedly annealed.

Figure 1 shows the temperature dependence of the a.c. resistance. The four-point measurement configuration is schematically presented in the inset. The current can be fed through the voltage contact area (a , dashed curves) or through the separate current contact (b). Curve b shows a sharp superconducting transition ($\leq 2 \text{ K}$) with a zero resistance state at 91 K. The contact resistance, the difference between the curves a and b , has a room-temperature value of $2.4 \text{ m}\Omega \text{ mm}^{-2}$ but decreases with decreasing temperature and drops to $< 10 \mu\Omega \text{ mm}^{-2}$ at 77 K. The highest value of the contact resistance measured in the dozens of contacts we have examined is $10^{-3} \Omega \text{ mm}^{-2}$.

In Fig. 2, we present the d.c. resistance measurements in the same configurations as Fig. 1 but now as a function of current density at 77 K. As one can see, above the critical current there is hardly any difference between the two curves a and b , which means that the resistance of the contact is practically negligible. Below the critical current the contact resistance essentially saturates at a low value of $\sim 10 \mu\Omega \text{ mm}^{-2}$. Our estimates of critical-current density presented in Fig. 2 are in the low limit due to the difficulty in determining the contact area between the grains. Our preliminary microscopic studies indicate that during the heat treatment at 900°C , the silver essentially fills the surface pores of granular Y-Ba-Cu-O, and improves the contact between the grains hence reducing the contact resistance. This makes it also feasible as a contacting method for films of $\geq 1 \mu\text{m}$ thickness⁵.

[‡] Permanent address: Institute of Solid State Physics of the Academy of Sciences of USSR, 142432, Chernogolovka, USSR

Received 2 July; accepted 21 July 1987.

1. Le Page, Y. *et al.* *Phys. Rev. B* (submitted).
2. Iye, Y., Tamegai, T., Takeya, H. & Takei, H. *Jap. J. appl. Phys.* (submitted).
3. Dinger, T. R., Worthington, T. K., Gallagher, W. J. & Sandstrom, R. L. *Phys. Rev. Lett.* **58**, 2687–2690 (1987).
4. Roth, R. S., Dennis, J. R. & Davis, K. L. *Adv. Ceramic Mat.* **2**, 303–312 (1987).
5. Hidaka, Y., Enomoto, Y., Suzuki, M., Oda, M. & Murakami, T. *Jap. J. appl. Phys.* **26**, L377–378 (1987).
6. Gallagher, P. K., O'Bryan, H. M., Jr, Sunshine, S. A. & Murphy, D. W. *Mater. Res. Bull.* **22**, 995–1006 (1987).
7. Tarascon, J. M., McKinnon, W. R., Greene, L. H., Hull, G. W. & Vogel, E. M. *Phys. Rev. B* (submitted).

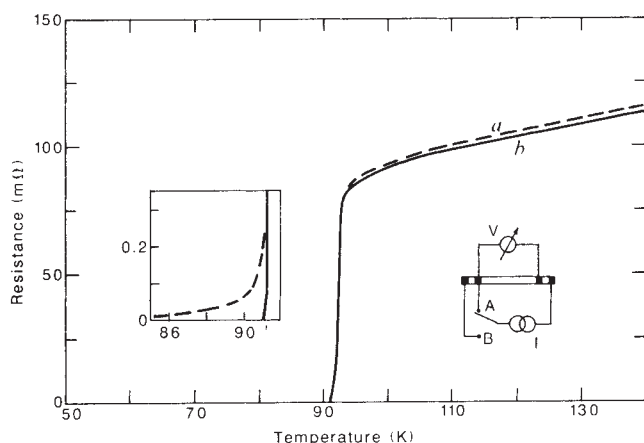


Fig. 1 Temperature dependence of the resistance of $\text{YBa}_2\text{Cu}_3\text{O}_7$ sample (curve *b*) and including a silver contact (dashed curve *a*). Right inset: the four-point measurement configurations *a* and *b*. Left inset: vertical scale expanded 100 times. The contact resistance has a room-temperature value of $2.4 \text{ m}\Omega \text{ mm}^{-2}$ and drops to less than $10 \text{ }\mu\Omega \text{ mm}^{-2}$ at 77 K. Sample size: $1 \times 1 \times 20 \text{ mm}^3$; current-contact area: 1 mm^2 ; $\rho(300 \text{ K}) = 1.2 \text{ m}\Omega \text{ cm}$; density $\approx 5.4 \text{ g cm}^{-3}$; $I = 6 \text{ mA}$. Sample densities in all experiments ranged from 3.5 to 5.4 g cm^{-3} .

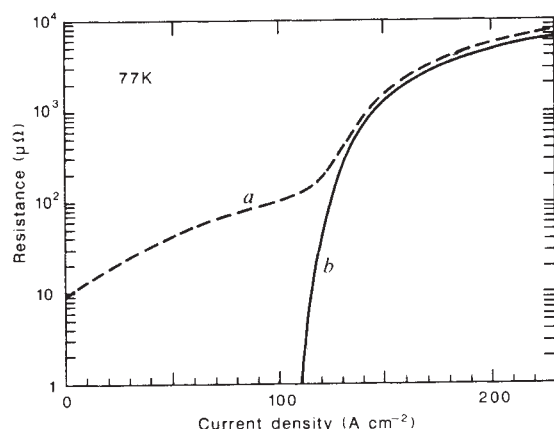


Fig. 2 The resistance of the sample alone (curve *b*) and including a silver contact (dashed curve *a*) as a function of current density in sample and contact at 77 K. Meaning of *a* and *b* as in Fig. 1. Below the critical current the contact resistance essentially saturates at the low value of $\sim 10 \text{ }\mu\Omega \text{ mm}^{-2}$.

Several groups discussed different types of contacts on copper-oxide superconductors including silver-paint and silver-epoxy⁶, platinum leads attached with gold-paste and pressed indium contacts⁷, and ultrasonic soldering⁸. We have tested several of these contacting techniques but the contact resistance is always $\sim 1 \text{ }\Omega \text{ mm}^{-2}$. On the other hand, whenever a silver coating has been baked onto the ceramic at 900°C , the contact resistance, and therefore the contact heating, were systematically reduced by several orders of magnitude. Indeed, through some of the contacts we passed more than $2,000 \text{ A cm}^{-2}$ (77 K) without burn-out. Scanning electron micrographs show that at the highest current densities it is the ceramic Y-Ba-Cu-O side of the contact area that burns out, rather than the mixed interface material of the contact itself.

Finally, we note that once the described silver contact has been prepared one can easily solder copper wires onto the contact area. But repeated soldering at 350°C increases the current dependence of the contact resistance, so that these soldering temperatures are better avoided.

We thank Professors R. Huguenin and L. Rinderer for stimulating discussions and Mr Burri for the microscopic investigations. We are grateful to the Swiss National Science Foundation for financial support, and the USSR Academy of Sciences for making this joint effort possible.

Received 1 July; accepted 10 July 1987.

1. Bednorz, J. G. & Mueller, K. A. *Z. Phys.* **B64**, 189-193 (1986).
2. Rao, C. N. R. *et al. Nature* **326**, 856-857 (1987).
3. Tholence, J.-L. & Raveau, B. *Nature* **327**, 658-659 (1987).
4. Cava, R. J. *et al. Phys. Rev. Lett.* **58**, 1676-1679 (1987).
5. Chaudhari, P. *et al. Phys. Rev. Lett.* **58**, 2684-2686 (1987).
6. Cava, R. J. *et al. Phys. Rev. Lett.* **58**, 408-410 (1987).
7. Chu, C. W. *et al. Phys. Rev. Lett.* **58**, 405-407 (1987).
8. Maeno, Y. *et al. Jap. J. appl. Phys.* **26**, L329-L331 (1987).

Incorporation of Pr in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$: electronic effects on superconductivity

L. Soderholm*, K. Zhang†, D. G. Hinks*, M. A. Beno*, J. D. Jorgensen*, C. U. Segre† & Ivan K. Schuller*

* Argonne National Laboratory, Argonne, Illinois 60439, USA

† Illinois Institute of Technology, Chicago, Illinois 60616, USA

Considerable excitement has been generated by the report of superconductivity above 90 K in a multi-phase Y-Ba-Cu oxide¹. The superconducting phase was subsequently identified as $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ ^{2,3}. The complete replacement of Y by the trivalent rare-earths La-Lu, with the exception of Ce, Pr and Tb, yields a superconducting phase with a critical temperature (T_c) almost identical to that reported for the yttrium compound⁴⁻⁷. This is somewhat surprising as most of these ions carry significant magnetic moments, the presence of which rapidly decreases T_c in ordinary superconductors. The three exceptions noted, Ce, Pr and Tb, are the rare-earth ions which have a stable tetravalent state. Here we present the results of a study of the structure and transport properties of the series of compounds $\text{Y}_{1-x}\text{Pr}_x\text{Ba}_2\text{Cu}_3\text{O}_{7-\delta}$ ($x = 0-1.0$), in which we observe a monotonic decrease in T_c and eventually a metal-insulator transition with increasing Pr concentration. To date, this is the only rare-earth ion which has been found substantially to alter T_c upon partial substitution for Y.

Samples were prepared by grinding and mixing Y_2O_3 , Pr_6O_{11} , BaCO_3 and CuO in the correct metal stoichiometries, and firing at 975°C , followed by annealing at 650°C and slow cooling in flowing oxygen. The samples were then reground and refired to improve homogeneity. Powder diffraction data were obtained using GE-XRD5 and SCINTAG-PADV X-ray diffractometers and the special-environment powder diffractometer (SEPD) at the intense pulsed neutron source (IPNS) at Argonne National Laboratory. Resistivities were measured in a closed cycle refrigerator from 300 to 8.5 K using Si-diode thermometry by the standard four-probe technique.

Diffraction measurements show this series of compounds to be predominantly a perovskite phase which is tripled along the *c*-direction. The majority phase in all samples can be roughly indexed according to the well-known orthorhombic *Pmmm* $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ structure^{8,9}. Although the structure can be indexed as orthorhombic through the full concentration range, the *a*- and *b*-axes converge with increasing *x*, indicating a tendency toward a tetragonal cell. The volume increases monotonically with Pr concentration consistent with solid solution formation. Preliminary Rietveld refinement of high-resolution neutron data indicates the presence of secondary phases ($<5\%$ of BaCuO_2 and CuO) together with a possible subtle structure change which complicates the analysis of oxygen site

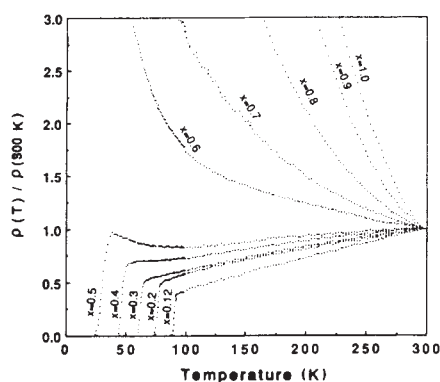


Fig. 1 Relative resistivity, ρ (normalized to room temperature) as a function of temperature for selected samples of $\text{Y}_{1-x}\text{Pr}_x\text{Ba}_2\text{Cu}_3\text{O}_{7-\delta}$. $x = 0.1$, $x = 0.2$, $x = 0.3$, $x = 0.4$, $x = 0.5$, $x = 0.6$, $x = 0.7$, $x = 0.8$, $x = 0.9$ and $x = 1.0$.

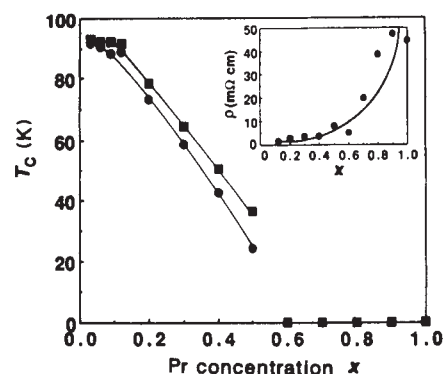


Fig. 2 Superconducting transition as a function of Pr concentration, x . (■) Indicates the onset in the resistivity while (●) shows the temperature at which the resistivity becomes zero. The inset shows the room temperature resistance as a function of Pr concentration.

occupancy. For this reason we have not included an analysis of the oxygen stoichiometry and its consequences.

Figure 1 shows relative resistivities as a function of temperature for selected Pr concentrations. The normal state resistivity exhibits a systematic metal (positive temperature coefficient of resistivity TCR) to non-metal (negative TCR) transition as a function of x for all samples. The superconducting transition temperature decreases monotonically, with samples above $x \approx 0.5$ becoming non-superconducting (Fig. 2). The absolute room-temperature resistance (inset Fig. 2) increases dramatically as the samples become non-metallic.

Changes in resistivity could occur if changes in grain sizes, special geometric arrangements of other phases, or general changes in microstructure are present. But while not studied as a function of x , the microstructure is not expected to vary in a fashion which could account for the observed systematic behaviour. Moreover, the fraction of minority phase present is always far from the three-dimensional percolation threshold.

The metal-insulator transition at $x \approx 0.5$ occurs at a much higher resistivity than commonly observed in most disordered systems¹⁰. This indicates that superconducting metal oxides behave in some sense as ordinary metals with a low carrier density, in contrast to a number of disordered two- and three-dimensional metals which show a universal metal-non-metal transition as a function of absolute resistivity (Mooij correlation)¹¹ for resistivities orders of magnitude lower ($\approx 150 \mu\Omega \text{ cm}$). Ordinarily the metal-insulator transition has been thought to occur close to the Ioffe-Reggel limit¹², where the electronic mean free path becomes comparable to the interatomic spacing. But in these materials, one might argue that the metal-insulator transition arises from the very low carrier density (assuming a constant mean free path), as the transition occurs when the resistivity is several orders of magnitude higher than the Ioffe-Reggel limit. By scaling the resistivity at which the metal-insulator transition occurs with that of simple metals one obtains a carrier density of 10^{20} – $10^{21} \text{ electrons cm}^{-3}$. The same kind of arguments about the resistivity have been invoked to claim low carrier density in $\text{LaNi}_{1-x}\text{M}_x\text{O}_3$ ($\text{M} = \text{Cr}, \text{Mn}, \text{Fe}$ and Co) perovskite systems¹³.

Similar trends in T_c and resistivity were observed on quenched samples of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ ^{14,15}, where the decrease in the transition temperature and resistance ratio along with an increase in the resistivity correlates with a decrease in oxygen stoichiometry. The lower oxygen content results in a decrease in the average oxidation state of Cu, with Cu^{3+} reduced to Cu^{2+} . As the behaviour of the resistivity and critical temperature of Pr-doped samples are much the same as that observed in quenched

samples, we postulate that doping with Pr is equivalent to decreasing the oxygen stoichiometry in undoped samples. With constant oxygen content the reduction of the formal charge state of Cu from Cu^{3+} to Cu^{2+} can be accounted for by incorporation of Pr in its tetravalent state. This hypothesis is supported by the increase in the unit cell volume with increasing Pr concentration, as the reduction of Cu results in an increase in the average Cu–O bond distance as electrons are added.

It was previously found that $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ undergoes a phase change from orthorhombic (with ordered Cu–O chains) to tetragonal (without the chains) when $\delta = 0.5$, as all the copper in the sample is reduced to Cu^{2+} (ref. 16). Equivalently, in the Pr samples $x = 0.6$ corresponds to the absence of Cu^{3+} (assuming $\delta = 0.2$), as each Pr^{4+} adds one electron to the structure. If the structural phase transition in the undoped samples is electronically driven, this might explain the possible subtle structural changes observed. Neutron diffraction experiments at high temperatures¹⁶ and on quenched samples (J. D. Jorgensen, unpublished data) show that the reduction in T_c , the trend toward a tetragonal structure¹⁴ and the decrease in oxygen stoichiometry are accompanied by a disordering of the sub-lattice containing the Cu–O chains. We therefore expect a similar disordering with increasing Pr concentration. It appears that the formal charge state of Cu and the oxygen vacancy ordering are intimately linked and therefore they both play a key role in the underlying superconducting mechanism in compounds which are isostructural with $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$.

We thank B. D. Dunlap, W. T. Carnall, J. Sullivan, M. B. Brodsky and K. Gray for helpful criticisms, and J.-P. Locquet for bringing ref. 13 to our attention. This work is supported by Basic Energy Sciences, Divisions of Chemical Sciences and Materials Sciences.

Received 30 June; accepted 10 July 1987.

- Wu, M. K. *et al.* *Phys. Rev. Lett.* **58**, 908 (1987).
- Hinks, D. G. *et al.* *Appl. Phys. Lett.* **50**, 1688 (1987).
- Tarascon, J. M. *et al.* *Phys. Rev. B* (submitted).
- Hor, P. H. *et al.* *Phys. Rev. Lett.* **58**, 18–19 (1987).
- Engler, E. M. *et al.* *J. Am. chem. Soc.* **109**, 2848 (1987).
- McKinnon, W. R. *et al.* *Proc. MRS Mtg* (1987).
- Tarascon, J. M. *et al.* *Proc. MRS Mtg* (1987).
- Beno, M. A. *et al.* *Appl. Phys. Lett.* **51**, 57 (1987).
- David, W. I. F. *et al.* *Nature* **327**, 310 (1987).
- Mott, N. *Metal-Insulator Transition* (Taylor-Francis, New York, 1974).
- Mooij, J. *Phys. Stat. Solid A17*, 521 (1973).
- Ioffe, A. & Reggel, A. *Prog. Semicond.* **4**, 237 (1960).
- Ganguly, P., Vasanthacharya, N. Y., Rao, C. N. R. & Edwards, P. P. *J. Solid State Chem.* **54**, 400 (1984).
- Schuller, Ivan K. *et al.* *Solid State Commun.* **63**, 385 (1987).
- van den Berg, J. *et al.* *Europhys Lett.* (submitted).
- Jorgensen, J. D. *et al.* *Phys. Rev. B* (in the press).

Crystal structure of the $\text{YBa}_2\text{Cu}_3\text{O}_7$ superconductor by single-crystal X-ray diffraction

G. Calestani & C. Rizzoli*

Istituto di Strutturistica Chimica, Università di Parma and *Centro di Studio per la Strutturistica Diffattometrica del CNR Viale delle Scienze, 43100 Parma, Italy

Just after Wu *et al.*¹ reported superconductivity above 90 K in barium, yttrium and copper ternary oxides, the formula ($\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$) and the structural arrangement (oxygen-deficient tripled-perovskite-type structure) of the superconducting phase was proposed by different groups. As theoretical proposals depend crucially on the details of the atomic arrangement reducing the dimensionality of the system, a number of structural determinations were carried out by X-ray, neutron and electron diffraction. Although these studies established the structure of the metal framework (the ordering of Ba and Y in the sequence Y–Ba–Ba being responsible for tripling the crystallographic c axis), discrepancies persist regarding the position of the oxygen vacancies. X-ray experiments on single crystals^{2–3} have reported that the oxygen vacancies are distributed in an ordered way in the plane containing Y atoms ($z = 1/2$) and statistically in the plane at $z = 0$. High-resolution neutron-powder diffraction^{4,5} showed, on the other hand, a reduction in the structure dimensionality. The oxygen vacancies in the plane at $z = 0$ are ordered along the a -axis, so that one-dimensional chains of square-coordinated copper ions parallel to the b – c plane are formed. These results, confirmed by electron diffraction⁶, justify the reduction of crystal symmetry from tetragonal to orthorhombic observed in several studies. Structural differences between X-ray and neutron diffraction studies were attributed⁴ to the fact that single crystals used in X-ray measurements were probably highly twinned. But neutron experiments need large amounts of material, increasing the possibility of the samples being inhomogeneous. Here we report X-ray structure determination on a $\text{YBa}_2\text{Cu}_3\text{O}_7$ single crystal showing a T_c above 90 K. The results, in agreement with neutron-powder diffraction, confirm the complete ordering of the oxygen vacancies leading to the one-dimensional chains.

Single crystals of $\text{YBa}_2\text{Cu}_3\text{O}_7$ were grown by reacting stoichiometric amounts of Y_2O_3 , CuO and BaO_2 . The powder was ground together and pressed into a pellet at 300 kg cm^{-2} , heated in a muffle furnace at 900°C in air for 8 h and finally cooled (50°C per h) in the furnace atmosphere. BaO_2 was used as the internal oxygen source. At the end of the treatment the samples appeared well crystallized, resulting in a block of almost equidimensional black prisms ranging from 50 to $200 \mu\text{m}$ on edge. Slow cooling was applied to facilitate the tetragonal-orthorhombic phase transition. It was shown⁷ by thermogravimetric study that thermal treatments in an air or oxygen

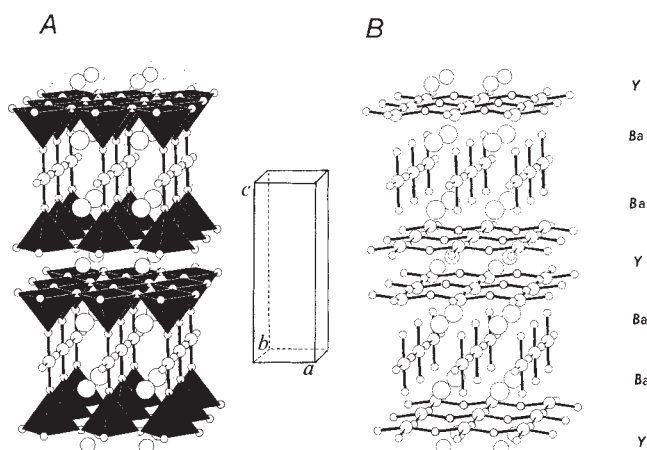


Fig. 1 Crystal structure representation of $\text{YBa}_2\text{Cu}_3\text{O}_7$ showing the copper coordinations (A) and the puckered sheets of square planar-like groups (B).

atmosphere are connected with a tetragonal-orthorhombic phase transition which plays an important part in determining the superconducting properties of the sample. The high-temperature tetragonal phase corresponds to the most reduced phase ($x > 0.3$) and could be related to superconductivity loss in samples annealed in air⁸ or quenched in an oxygen atmosphere. The orthorhombic phase which appears at $x < 0.3$ seems to be responsible for superconductivity. The highest T_c and sharpest transitions were detected for samples close to the O_7 composition ($x \approx 0$) and showing orthorhombic distortion of the perovskite structure. A low-field microwave absorption technique proposed by Blazey *et al.*⁹ was used for testing superconductivity in selected single crystals and proved to be a powerful and simple method for characterizing these small samples. The non-resonant microwave absorption detected by liquid-nitrogen cooling revealed a T_c above 90 K and a surprisingly sharp transition (width $< 1 \text{ K}$).

Even if larger crystals were available, a well-formed single crystal of $\sim 90 \times 110 \times 110 \mu\text{m}^3$ was chosen for final data collection, its dimensions representing a compromise between reliable diffraction intensities and minimization of absorption effects. Diffraction data were collected at room temperature on a Philips PW 1100 single-crystal diffractometer using graphite monochromatized $\text{Mo K}\alpha$ radiation. Lattice parameters $a = 3.827(1)$, $b = 3.893(1)$ and $c = 11.699(2) \text{ \AA}$ were refined by least-squares fitting of the setting angles of 20 selected reflections in the 2θ range 14 – 24° . We measured 347 reflections in the range $3 < \theta < 30$, of which 217 showing integrated intensities $I > 2\sigma(I)$ corrected for Lorentz and polarization effects were used in the structure analysis. According to systematic absences the space group $Pmmm$ was chosen for structure refinement. Starting from the atomic coordinates of the cations determined in previous studies^{2–5}, the oxygen atoms were located by difference Fourier

Table 1 Fractional atomic coordinates and thermal parameters ($\times 10^4 \text{ \AA}^2$) for $\text{YBa}_2\text{Cu}_3\text{O}_7$

Atom	a/a	y/b	z/c	U_{11}	U_{22}	U_{33}
Y	0.5000	0.5000	0.5000	123 (26)	162 (29)	1 (30)
Ba	0.5000	0.5000	0.1850 (2)	68 (10)	45 (9)	87 (14)
Cu (1)	0.0000	0.0000	0.0000	144 (36)	237 (44)	32 (33)
Cu (2)	0.0000	0.0000	0.3565 (5)	55 (21)	57 (20)	90 (32)
O (1)	0.0000	0.5000	0.0000	587 (360)	68 (196)	35 (158)
O (2)	0.0000	0.0000	0.1566 (23)	1 (98)	226 (130)	59 (89)
O (3)	0.5000	0.0000	0.3776 (21)	23 (106)	112 (125)	168 (121)
O (4)	0.0000	0.5000	0.3765 (21)	22 (113)	256 (172)	131 (122)

Thermal parameters are in the form: $\exp(-2\pi^2 U_{11} h^2 a^{*2} + \dots + U_{12} hka^*b^*)$.
 $U_{12} = U_{23} = U_{13} = 0$ by symmetry requirements.

maps on four sites of the lattice according to the neutron-powder diffraction experiments^{4,5}. Refinement of the oxygen site occupancies converged to one location with an error within three standard deviations, showing a stoichiometry close to $\text{YBa}_2\text{Cu}_3\text{O}_7$. They were therefore blocked in the final calculations. After absorption corrections were performed with the program ABSORB¹⁰ (with a correction range in angles ϕ and μ of 1.30–0.67), refinement cycles with anisotropic thermal parameters for all the atoms lead to a final $R = 0.054$ based on observed and calculated factors ($R = \sum |kF_o - |F_c|| / \sum kF_o$) for 217 unique reflections. Introduction of the alternative space group $Pm2m$ suggested by Beyers *et al.*¹¹ on the basis of electron diffraction data showed no evidence of lacking a centre of symmetry. All calculations were performed on an IBM-PC AT with SHELX-76¹² integrated in the CRYSRULER package¹³.

Table 1 lists the final atomic coordinates and thermal parameters. Positional parameters are comparable with those found by powder-neutron diffraction^{4,5}. Table 2 reports selected bond

Table 2 Selected bond distances (Å) and angles (°) for $\text{YBa}_2\text{Cu}_3\text{O}_7$

Y–O (3)	4 × 2.418 (15)	Cu (1)–O (1)	2 × 1.947 (5)
Y–O (4)	4 × 2.399 (15)	Cu (1)–O (2)	2 × 1.834 (27)
		Cu (2)–O (2)	2.341 (28)
Ba–O (1)	2 × 2.891 (2)	Cu (2)–O (3)	2 × 1.929 (3)
Ba–O (2)	4 × 2.750 (3)	Cu (2)–O (4)	2 × 1.961 (3)
Ba–O (3)	2 × 2.980 (19)		
Ba–O (4)	2 × 2.948 (19)		
O (1)–Cu (1)–O (2)	90.0 (5)		
O (2)–Cu (1)–O (2)	180.0 (7)		
O (1)–Cu (1)–O (1)	180.0 (0)		
O (4)–Cu (2)–O (4)	166.3 (1)		
O (3)–Cu (2)–O (3)	165.3 (7)		
O (3)–Cu (2)–O (4)	89.1 (0)		
O (2)–Cu (2)–O (4)	96.9 (6)		
O (2)–Cu (2)–O (3)	97.4 (6)		

The numbers in parentheses are estimated standard deviations.

distances and angles. Figure 1 shows the crystal structure of $\text{YBa}_2\text{Cu}_3\text{O}_7$. Oxygen vacancies are localized in an ordered way at (0, 0, 1/2) (in the plane containing Y atoms) and at (1/2, 0, 0) (in the plane containing the Cu(1) atoms), producing two crystallographically and chemically different copper sites. Cu(1) sites are surrounded by square planar oxygen coordinations which are linked together in the b - c plane forming one-dimensional chains parallel to the b -axis. As reported by David *et al.*⁵ the unexpectedly short Cu(1)–O(2) bond distance suggests, on the basis of bond strength calculations, that this site is preferentially occupied by the Cu^{3+} ions present in the structure. Cu(2) is fivefold coordinated in a square pyramid arrangement of oxygen atoms. The planar copper–oxygen bonds range between 1.93 and 1.96 Å and are significantly shorter than the apical one (Cu(2)–O(2) = 2.34 Å). The structure can therefore be regarded as consisting of puckered two-dimensional sheets of CuO_4 planar-like groups, corner-linked in the a - b plane, derived from the square pyramids which interact weakly with the infinite one-dimensional chains (Fig. 1B). This arrangement is reflected in the large anisotropy of the thermal parameters of the oxygen atoms. O(1) shows a large component of the thermal motion along the a -axis, indicating large vibrational movement of the chains, possibly due to the lack of linking amongst the one-dimensional chains in the a -direction. O(2) owns a large component along the b -axis, while the thermal motion of O(3) and O(4) develops in the b - c plane.

Due to the presence of the one-dimensional chains, no phase transition to tetragonal symmetry is possible without destroying the typical features of the structure. The distribution of the disordered vacancies found in X-ray single-crystal structure determination could be associated with the tetragonal reduced phase or with twinning effects. The agreement of the present

structure determination performed on a single crystal showing superconductivity above 90 K with those resulting from neutron and electron diffraction should confirm that the symmetry reduction observed in the superconducting phase is unambiguously related to an ordering of the oxygen vacancies which is fundamental for the understanding of the mechanism of superconductivity in this compound.

We thank G. Amoretti, E. Buluggiu, D. C. Giori and A. Vera for the EPR measurements.

Received 13 July; accepted 17 July 1987.

1. Wu, M. K. *et al. Phys. Rev. Lett.* **58**, 908–910 (1987).
2. Hazen, R. M. *et al. Phys. Rev. B* **35**, 7228–7241 (1987).
3. Siegrist, T. *et al. Phys. Rev. B* **35**, 7137–7139 (1987).
4. Capponi, J. J. *et al. Europhys. Lett.* **3**, 1301–1307 (1987).
5. David, W. I. F. *et al. Nature* **327**, 310–312 (1987).
6. Ourmazd, A. *et al. Nature* **327**, 308–310 (1987).
7. Strobel, P. *et al. Nature* **327**, 306–308 (1987).
8. Gallager, W. J., Sandstrom, R. L., Dinger, T. R., Shaw, T. M. & Chance, D. A. *Solid State Commun.* (submitted).
9. Blazey, K. W. *et al. Phys. Rev. Lett.* (submitted).
10. Ugozzoli, F. *Comput. Chem.* **11**, 109–120 (1987).
11. Beyers, R. *et al. Appl. Phys. Lett.* (submitted).
12. Sheldrick, G. M. *SHELX-76* (University of Cambridge, Cambridge, 1976).
13. Rizzoli, C., Sangermano, V., Calestani, G. & Andreotti, G. D. *J. appl. Cryst.* **20**, (in the press).

Meteoroid sonic shock-wave-generated seismic signals observed at a seismic array

F. M. Anglin & R. A. W. Haddon

Geophysical Division, Geological Survey of Canada,
1 Observatory Crescent, Ottawa, Ontario, K1A 0Y3, Canada

At approximately 04:20 local time (10:20 UT) on 19 September 1986, a local resident observed a meteor in the sky above Yellowknife, North-west Territory, Canada. We subsequently found seismic signals, recorded at the Yellowknife seismic array in North-west Territory, Canada, that coincided with the visual observation of the meteor in the sky overhead. The signals cannot be explained by any earthquake, explosion, or impact-generating mechanism. They are shown to be consistent with their having been generated by the shock wave created by the associated meteoroid as it passed through the atmosphere overhead. The meteoroid's trajectory and velocity, as inferred from the array records, is consistent both with the observed direction of travel of the meteoroid and independent seismic observations recorded at two single stations in the Northern Canadian seismic network.

The seismic data were sent to the Geophysics Division of the Geological Survey of Canada in Ottawa for analysis with a view to locating the point of presumed impact with the Earth. Initial scrutiny of the data revealed severe paradoxes on any conventional interpretation for a point-source impact. If the data had resulted either from the impact of the meteoroid with the Earth or from a natural earthquake or from an explosion, the lack of any measurable curvature in the relative arrival times on the west-east R line of Fig. 1 would imply that the source must be located at a distance of at least 250 km from the array. But the pulse-like simplicity of the main observed waveforms is inconsistent with all experience for a source at this distance. This and other such paradoxes led us to consider, as an alternative interpretation, that the observed seismic signals had been generated by the shock wave created by the meteoroid as it passed through the atmosphere overhead.

The Yellowknife Array has been described by Manchee and Somers¹ and consists of two orthogonal lines each containing nine equally spaced (2.5-km apart) short-period vertical seismometers. One of the lines (B2 to B10) is set in the south-north

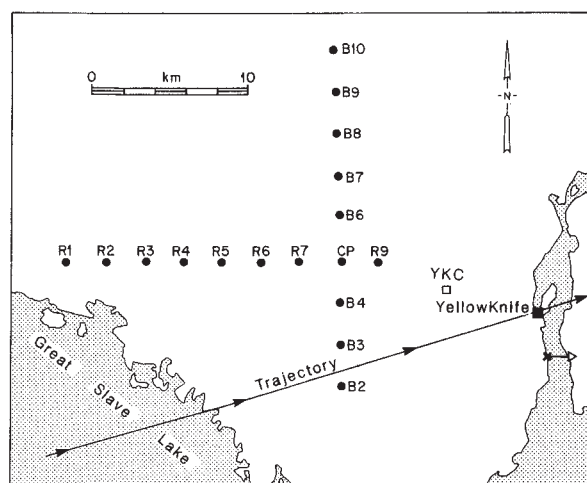


Fig. 1 Yellowknife Array configuration with short-period vertical seismometers at R1 to R9 and B2 to B10 and the Standard Seismological Station at YKC. The point of visual observation and the approximate observed direction of travel of the meteoroid are shown at X. The path inferred from the array data is shown by the line passing between B2 and B3.

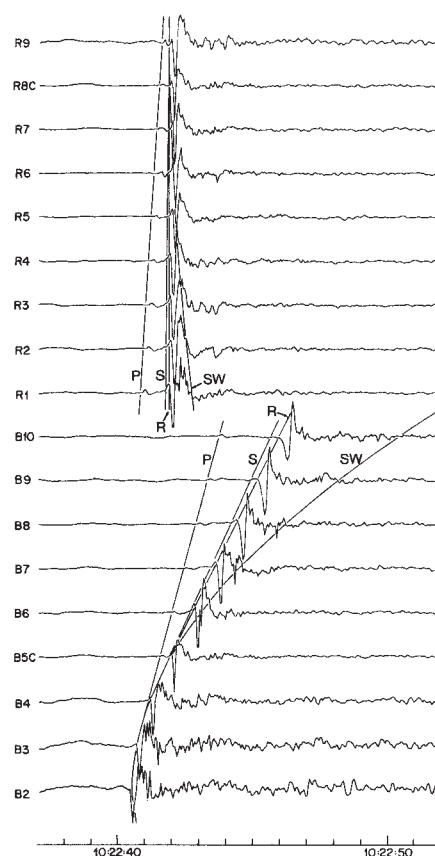


Fig. 2 The seismic signals recorded on the 17 channels of array data (R8 and B5 are a common centre-point channel) with UT time in hours, minutes and seconds (UT) indicated. The superimposed curves are the theoretical travel-time curves for a meteoroid source at an elevation of 45 km travelling at a velocity of 12 km s^{-1} at 15° to the north of east. The seismic propagation velocities taken for P, S and Rayleigh waves were 6.4, 3.55 and 3.3 km s^{-1} and the velocity of the shock wave was assumed to be 0.3 km s^{-1} (which is the approximate velocity of sound in air). As described in the text, a range of different solutions consistent with the data can be obtained for larger assumed shock-wave velocities. The nominal maximum peak to peak amplitudes are $1,000 \text{ nm s}^{-1}$ for B5 to B9 and R1 to R8 and 500 nm s^{-1} for B2 to B4 and R9.

direction and the other (R1 to R9) in the west-east direction. The array configuration is shown in Fig. 1.

The seismograms in Fig. 2 show that the main signals observed on the R1 to R9 instruments arrived almost simultaneously, while on the B2 to B10 instruments they arrived with a significantly nonlinear step-out, the earliest arriving on B2 and the latest on B10. Small precursory signals showing a nearly linear step-out can also be clearly identified in addition to the main signals in each of the records from R1 to R9 and B5 to B10.

To investigate the possibility that the signals had been generated by the shock wave created by the meteoroid, we took as a trial hypothesis the idea that the path of the meteoroid was horizontal and that the speed of the meteoroid was large compared with the speed of the shock wave. In these circumstances the wavefront of the shock wave is a cone of narrow vertex angle² and at any instant the line of intercept of the wavefront with the Earth's surface would be a hyperbolic curve with its axis parallel to the meteoroid's path. At points suitably distant from the vertex of the hyperbola, the wavefront can be visualized as approximately cylindrical. The speed at which it sweeps across the Earth's surface decreases monotonically from extremely high values at points nearly vertically below the meteoroid's path, to reach the value of the speed of the shock wave as the radius of the approximately cylindrical wavefront increases.

Wave theory calculations, as described by Ewing *et al.*, show that a cylindrical wave impinging upon the Earth's surface would generate P, S and Rayleigh waves. At suitably distant points the arrival times of each are precisely the same as if each were instantaneously generated when and where the sweep velocity of the line of intercept of the shock wavefront with the Earth's surface is equal to the propagation velocity of that particular wave.

The theoretical travel-time curves shown in Fig. 2 have been computed on this basis. For the particular curves shown the meteoroid's velocity and elevation are taken as 12 km s^{-1} and 45 km, respectively, and its path is taken as horizontal and passing vertically over a point midway between instruments B2 and B3 in a direction 15° north of east as shown in Fig. 1. The theoretical travel-time curves shown in Fig. 2 labelled P, S, R and SW correspond to the P, S and Rayleigh waves and the shock wavefront, respectively. The qualitative features of the waveforms are completely explained by the theoretical curves

and the observed times are matched to within $\pm 0.1 \text{ s}$. In obtaining the theoretical results shown in Fig. 2 we assumed that the shock wave had decayed into a sound wave (Mach wave) with a propagation velocity of 0.3 km s^{-1} , before reaching the Earth's surface. The referee, P. D. Marshall, queried this assumption and suggested that it was probably still a shock wave. If so, it could have a much higher velocity than the velocity assumed, the actual velocity depending critically on the amplitude and decreasing as the amplitude decays. He also pointed out: (1) that the observed direction of first motion of the observed Rayleigh wave is indicative of a compressional source loading of the Earth's surface; (2) that a strong shock can generate Rayleigh waves of very large amplitude relative to the P-wave amplitude as a result of resonant coupling; and (3) that this phenomenon is highly diagnostic of atmospheric explosions.

On our original assumptions we had estimated that the uncertainties in the theoretical results shown in Fig. 2 were about $\pm 10 \text{ km}$, $\pm 5 \text{ km s}^{-1}$, $\pm 3^\circ$ and $\pm 3^\circ$ for the elevation, velocity, direction and inclination of the meteoroid's path, respectively. In the light of the referee's comments we have derived further models for various assumed shock-wave propagation velocities from 0.3 to 3.0 km s^{-1} and have obtained equally good fits to the data for all velocities in this range. The main effects of

increasing the assumed shock-wave velocity are that the elevation of the meteoroid must be correspondingly reduced and the curve marked SW moves closer to the curves marked S and R in Fig. 2. For the largest shock-wave velocity assumed (3.0 km s^{-1}), the values obtained for the elevation, velocity, direction and inclination are $(4 \pm 1) \text{ km}$, $(15 \pm 5) \text{ km s}^{-1}$ ($11 \pm 3^\circ$) and $(0 \pm 3^\circ)$. Thus, while the inferred elevation of the meteoroid is critically dependent on the shock-wave velocity, the other inferred parameters of the trajectory are not. In the absence of any independent evidence on the characteristics of the shock wave, little more can be said.

Supporting our interpretation, we learned after obtaining the above results that the meteoroid was observed to disappear over the horizon in the east (Mrs Laurie Harder, personal communication). We have also searched the records of those seismic stations in the Northern Canadian Seismic Network located in the vicinity of the inferred path and found that sharp signals were recorded at the Yellowknife standard station (YKA) at 10:22:41.5 UT and at Fort Simpson (FST) (which is 350 km west of Yellowknife) at 10:20:19.1 UT. The fact that no other signals were observed at either of these stations within at least an hour of the above times and that the latter are fully consistent with the results inferred from the array data adds further strong support.

Folinsbee *et al.*⁴ have previously reported certain seismic signals recorded on a three-component short-period seismogram which they have interpreted as resulting from the direct sonic wave of a meteoroid while Halliday and Blackwell⁵ have previously reported seismic observations which they interpret as resulting from a meteoroid impact with the ground. To our knowledge the observations reported in this letter provide the first unambiguous observations of a shock wave and associated seismic waves generated by a meteoroid and observed at an array of closely spaced seismometers.

We thank Mr D. Monsees and the Yellowknife station staff for the data from Yellowknife, Mr J. A. Haddon for computing assistance and Dr R. North and Mr R. Wetmiller for helpful suggestions on the manuscript. We also gratefully acknowledge Dr P. D. Marshall's valuable contributions.

Received 13 February; accepted 28 May 1987.

1. Manchee, E. B. & Somers, H. *Publs dom. Obs.* **32**, 67–84 (1966).
2. Von Karman, T. *J. aeronaut. Sci.* **14**, 373–402 (1947).
3. Ewing, W. M., Jardetski, W. S. & Press, F. *Elastic Waves in Layered Media* (McGraw-Hill, New York, 1967).
4. Folinsbee, R. E., Bayrock, L. A., Cumming, G. L. & Smith, D. G. W. *J. R. astr. Soc. Can.* **63**, 61–86 (1969).
5. Halliday, I. & Blackwell, A. T. *Meteoritics* **6**, 39–47 (1971).

Cold climatic episode of Younger Dryas age in Tierra del Fuego

C. J. Heusser* & J. Rabassa†

* Department of Biology, New York University, 1009 Main, New York 10003, USA

† Centro Austral de Investigaciones Científicas, C. C. 92, 9410 Ushuaia, Tierra del Fuego, Argentina

Pronounced climatic change during the last glacial maximum is recognized as global in extent, and has been attributed to orbital forcing¹. Not readily explained in a global context because of the equivocal nature of the evidence are the late-glacial climatic events of short (100–1,000-yr) duration. The cold climate of the Younger Dryas chron (~11,000–10,000 yr BP), for example, is known to have been particularly pronounced over the North Atlantic and parts of Europe and North America^{2,3}, whereas in the Southern Hemisphere, cold conditions for that time are not well established. Here we present evidence for late-glacial climatic cooling in southernmost South America from studies of radiocarbon-dated glacial and pollen stratigraphy at sites along the Beagle Channel in Tierra del Fuego, where a forest–tundra tension zone in the region is especially sensitive to shifting climate. We conclude that the cold episode corresponds with the time of the Younger Dryas in Europe, when a drop in temperature, ascribed most recently to colder sea surface temperatures and production rate of deep water in the North Atlantic^{4,5}, apparently occurred worldwide.

Beagle Channel (~55°S) is a drowned, east–west-trending, glacial valley lying along the southern edge of Isla Grande de Tierra del Fuego (Fig. 1). During the last glacial maximum, dated at between 20,000 and 18,000 yr BP in southern South America⁶, a glacier flowed ~200 km eastward in Beagle Channel from an ice cap centred in the Cordillera Darwin to reach its farthest point at Isla Gable, according to Caldenius⁷, where it formed a prominent end moraine. The Cordillera Darwin, today occupied by a complex of ice fields and glaciers, is at altitudes of up to almost 2,500 m. During the Pleistocene it was the primary source for regional glaciers draining to the Pacific Ocean and, by way of the Strait of Magellan, to the Atlantic.

No bracketing ages yet apply to the emplacement of the Isla Gable moraine, leaving open the possibility that it represents a

younger re-advance of 14,500–14,000 yr BP⁶. The Beagle Channel glacier receded from the moraine before 12,730 ± 90 yr BP (QL-1685), the age of near-basal sediments in a topogenous mire at Caleta Róbalo located just west of Puerto Williams and ~7 km inside the moraine limit (Fig. 1). Later, under conditions of rising sea-level^{8,9}, the ice front withdrew to the west to form the morainal complex at Península Ushuaia with erosional remnants of drift on nearby islands. At Punta Pinguinos on Península Ushuaia, deposits exposed in the sea cliff consist of lodgment till, above and below tide level, overlain by ice-contact stratified drift and, in turn, by peat interbedded with pebbly sand, laminated clay and gravel. A date of 10,080 ± 280 yr BP (RL-1998) for the lower ± 5 mm of peat in contact with the drift provides a minimum age for ice recession. This age for deglaciated ground locally is corroborated by a bottom date of 10,080 ± 250 yr BP (RL-2001) from a raised mire at Lapataia, 30 km west of Punta Pinguinos. Thus the formation of the Península Ushuaia morainal complex is a late-glacial event dating from after 12,730 and before 10,080 yr BP.

Pollen analysis of the radiocarbon-dated mire section at Caleta Róbalo (Figs 2, 3) provides vegetational and climatic background and refinement of the time the Península Ushuaia moraine was emplaced. The site of the section, at present situated in Magellanic deciduous forest (Fig. 1) of beech (*Nothofagus pumilio* and *N. antarctica*)^{10,11}, was apparently located in tundra for ~3,000 yr in the late-glacial record. Pollen frequency during this interval (pollen assemblage zones 3a and 3b, Fig. 3) shows a decrease of *Empetrum* and increase of Gramineae, *Acaena* and Umbelliferae in zone 3a, implying expansion of herb tundra at the expense of dwarf shrub tundra (alpine meadow and dwarf-shrub heath type vegetation of Moore¹⁰), caused by a shift to colder climate.

Late-glacial changes are illustrated in Fig. 2 by pollen/spore influx, substantiating the trends in pollen frequency set up by shrub and herb taxa. The sharp drop in influx (overall, 387 to 57 grains cm⁻² yr⁻¹) after 11,850 ± 50 yr BP (QL-1729) with low values lasting until at least 10,510 ± 80 yr BP (QL-1684) in pollen assemblage zone 3a (830–790 cm in the section) suggests sparseness of tundra effected by a cold climate. Also indicative of cooling is the diminution in *Nothofagus* influx (860–830 cm) preceding levels at which all taxa reach minimum values. Low influx before 12,730 yr BP also implies an earlier interval of cold. In zone 3a, after 10,510 yr BP, *Nothofagus* increases in proximity

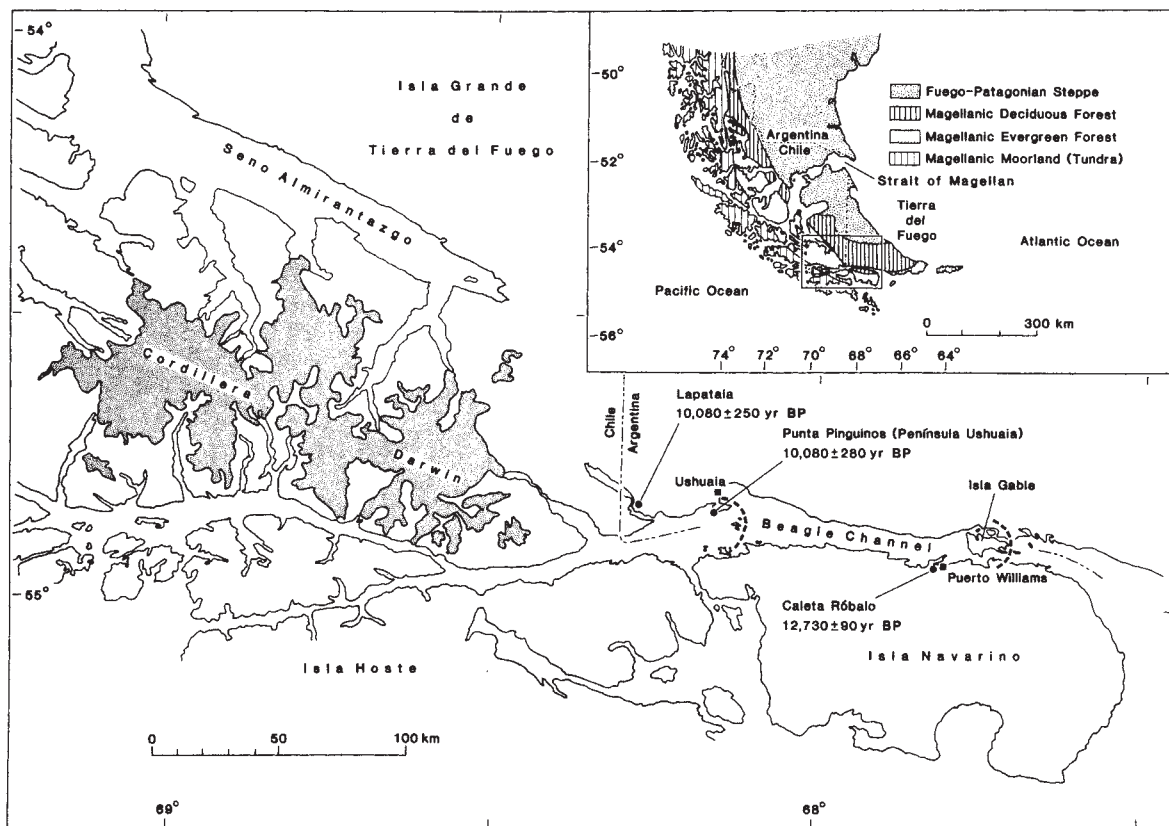


Fig. 1 Location of radioacarbon-dated sites in relation to end moraines (dashed arcuate lines, shown approximately) at Isla Gable and Península Ushuaia which formed in Beagle Channel by a late Quaternary glacier originating in the Cordillera Darwin (stippled areas are present-day icefields and glaciers). Inset shows modern distribution of vegetation formations^{15,16} and location of Beagle Channel region.

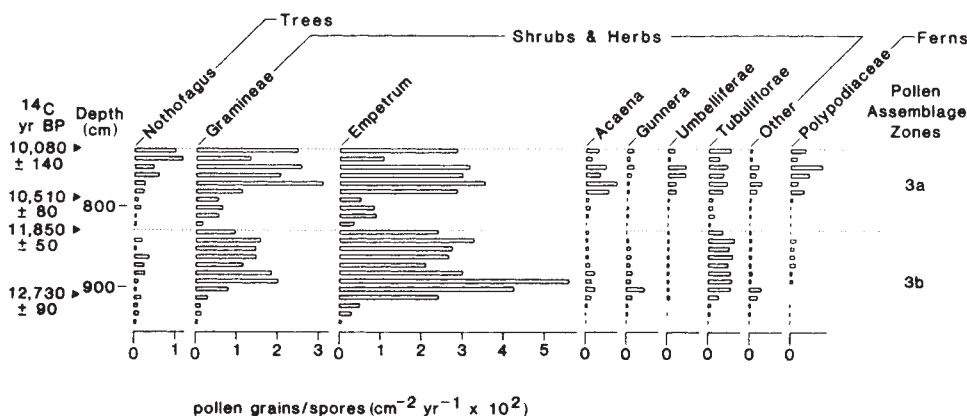


Fig. 2 Pollen and spore influx (number cm⁻² yr⁻¹) diagram of late-glacial pollen assemblage zones 3a and 3b (Fig. 2). Low influx values after 11,850 until 10,510 yr BP (zone 3a), indicative of an impoverished tundra resulting from cold climate, imply the time of glacial re-advance in Beagle Channel.

to the late-glacial-postglacial boundary, dated 10,080 ± 140 yr BP (QL-1718), suggesting early expansion of beech which gained dominance in zone 1b during the late postglacial (Fig. 3).

These data place late-glacial moraine formation at Península Ushuaia after 11,850 and until at least 10,510 yr BP during the interval of low influx of pollen/spores. They constitute the first pollen stratigraphic evidence together with dated glacial stratigraphy from a single region to support a re-advance of ice in southern South America in the time frame of the European Younger Dryas chron (~11,000–10,000 yr BP).

Correlative pollen records (refs 12–15) from the lake district of Chile (39°–41° 20' S) have been disputed because glacial moraines of Younger Dryas age were not reported. After full-glacial conditions 20,000–18,000 yr BP, glaciers in southern Chile and Argentina re-advanced over the lowland 14,500–14,000 yr BP during the closing millennia of the last ice age,

after which they were believed to have pulled back into their cordilleran source regions without advancing again until the onset of Neoglacial time^{6,16}. Fossil beetle evidence from the Chilean lake district^{17,18}, implying moderating climate after 14,000 yr BP, has also challenged the implication of pollen data for an episode of cold climate and glacial re-advance at the end of the late-glacial. Note, however, that the glacial history of this region comes from the lowland, whereas the record of glaciation in the interior of the nearby Andes Mountains is almost unknown. Unpublished data by C.J.H. indicate deglaciation before 12,300 yr BP at higher elevations in the Andes, where moraines of Younger Dryas age would have been laid down. The disparity between interpretations of pollen and beetle assemblage data is perhaps a reflection of site location and of differential sensitivity and response of taxa to ecological thresholds². Assemblages from ecotones are expected to be particularly

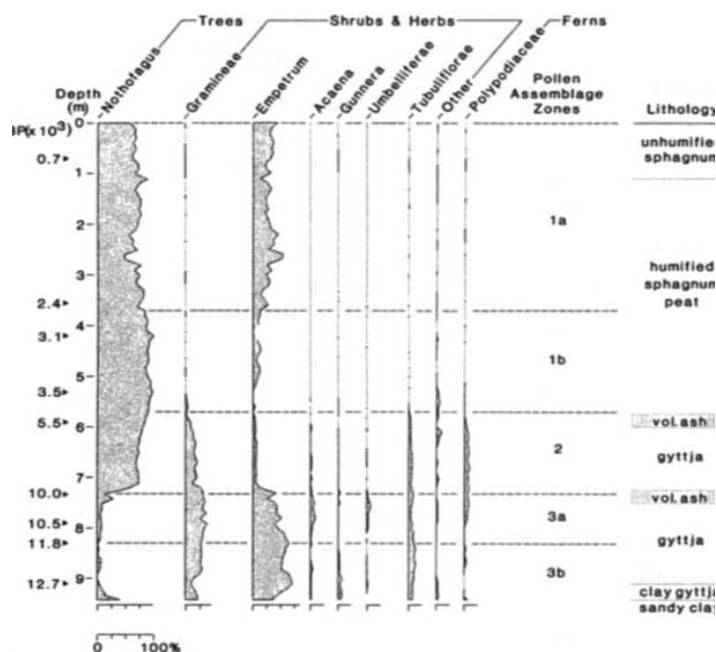


Fig. 3 Consolidated frequency pollen and spore diagram of section of mire at Caleta Róbal. Prevailing shrubs and herbs (Gramineae and *Empetrum* as dominants) show tundra during the first 3,000 yr (pollen assemblage zones 3a and 3b) covering the interval during which a glacial re-advance apparently occurred in Beagle Channel. Trees (*Nothofagus*) predominate after approximately 10,000 yr BP, indicating with Gramineae (zone 2) community openness at first, followed by closed forest (zone 1b), and later, by the spread of *Empetrum* (zone 1a), heath and mired ground. The profile of other shrub and herb taxa consists of *Berberis*, *Chenopodiaceae*, *Caryophyllaceae*, *Armeria*, *Misodendrum*, *Maytenus*, *Plantago*, *Rubiaceae* and *Liguliflorae*.

sensitive in this regard. The preservation of fossil assemblages transported beyond their natural environmental setting is also of interpretative significance. Pellets containing exoskeletons of woodland beetles are occasionally unearthed from treeless bogs where they have been carried in by birds and regurgitated¹⁹.

Despite this "intriguing enigma"²⁰, which may derive in part from differences in the indicator value of biota, elsewhere in the Southern Hemisphere there is evidence for climatic fluctuation at the close of the late glacial. At lower latitudes of South America, data from Colombia²¹, Ecuador²² and Peru²³ indicate an episode of colder climate and glacial re-advance between ~12,000 and 10,000 yr BP. In the El Abra corridor of Colombia, dates for this interval, the El Abra Stadial, place it more precisely

at 11,000–10,000 yr BP^{24,25}. At higher latitudes of the Hemisphere, problems of dating and interpretative timing of recognized late-glacial events in Antarctica Georgia³⁰ and New Zealand^{31,32}. Measurements of the Byrd, Dome C and Vostok ice cores from Antarctica consistently show what is interpretable as a late-glacial. Because of the proximal geographic relation to the dated Younger Dryas record in Tierra del Fuego to sites, correlation is strongly suggested. Biostratigraphy from marine cores taken in the Southern Ocean up to do not show the Younger Dryas event, which in this time frame is apparently beyond the limit of high-resolution sampling techniques (J. J. Morley, personal communication).

Evidence for the Younger Dryas is scattered in the hemispheres with site concentration in regions of the Atlantic Ocean⁴, where cold climate intensified, as by vegetation response² and glacier re-advance in northern Europe^{33,34} and by modification of vegetation in the Provinces of Canada³. Over extensive portions of the data are equally equivocal or wanting, indicating not only affected by the change of climate apparently by altitudinal–latitudinal constraints, differential lag of response to air–sea interaction, biological boundaries and time lag of climatically sensitive biota, among other factors. The unevenness in the evidence is perhaps in keeping with the response of biological systems to short-term climatic variability. For example, during the Little Ice Age centuries, which was on a shorter timescale than the Younger Dryas, glacial moraines indicative of climate fluctuations are found in mountain systems over the globe. The pollen data by and large fail to record the event.

Correspondence to the European Younger Dryas episode implied by the pollen and glacial stratigraphy from Tierra del Fuego suggests a response to conditions to both polar hemispheres. Propositions set forth to the climatic requirement involve, for example, regimes of surface temperatures⁴, production rates of deep-water and natural carbon dioxide variations in high-latitude. Floating ice reaching contiguous borderland and contributing to regional intensification of cold climate, as proposed for the North Atlantic³⁶, also may have been effective in the Southern Ocean. Tierra del Fuego at the southern tip of South America reaches farthest poleward of any major land mass in the Southern Hemisphere. Any extension of the late-glacial surrounding Antarctica may therefore result in a cooling of the Fuegian archipelago.

This study was supported by the NSF and by grant from Consejo Nacional de Investigaciones Científicas y Argentina. We thank L. E. Heusser, D. H. Rind and Peteet for critical comments.

Received 3 February; accepted 8 May 1987.

- Hays, J., Imbrie, J. & Shackleton, N. *Science* **194**, 1121–1132 (1981).
- Watts, W. A. in *Studies in the Late-Glacial of North-West Europe* (eds Lowe, J. J., Gray, J. M. & Robinson, J. E.) 1–21 (Pergamon, Oxford, 1979).
- Mott, R. J., Grant, D. R., Stea, R. & Occhietti, S. *Nature* **323**, 247–250 (1986).
- Rind, D., Peteet, D., Broecker, W., McIntyre, A. & Ruddiman, W. *Climate Dynam.* **1**, 3–33 (1986).
- Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21–25 (1985).
- Mercer, J. H. A. *Rev. Earth planet. Sci.* **11**, 99–132 (1983).
- Caldenius, C. C. *Geogr. Annl.* **14**, 1–164 (1932).
- Porter, S. C., Stuiver, M. & Heusser, C. J. *Quat. Res.* **22**, 59–67 (1984).
- Rabassa, J., Heusser, C. & Stuckenrath, R. *Quat. S. Am. Antarct. Pen.* (in the press).
- Moore, D. M. *Flora of Tierra del Fuego* 30–31 (Anthony Nelson, Oswestry, 1983).
- Pisano V., E. *Am. Inst. Patagonia* **12**, 159–171 (1981).
- Heusser, C. J. *Proc. Am. phil. Soc.* **110**, 269–305 (1966).

- Coles, J. M. & Orme, B. J. *Prehistory of the Somerset Levels* (Austin & Sons, 1975).
- Clapperton, C. M. *Quat. Sci. Rev.* **2**, 83–155 (1983).
- Hammen, T. van der, Barends, J., De Jong, H. & De Veer, A. A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **32**, 247–340 (1980/1981).
- Clapperton, C. M. *Quat. S. Am. Antarct. Pen.* **3**, 149–158 (1985).
- Wright, H. E. Jr *Quat. Res.* **21**, 275–285 (1984).
- Hammen, T. van der *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **25**, 111–164 (1980/1981).
- Schreve-Brinkman, E. J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **25**, 1–11 (1980/1981).
- Clapperton, C. M. & Sugden, D. E. *Quat. Res.* **18**, 243–267 (1982).
- Johnsen, S. J., Dansgaard, W., Clausen, H. B. & Langway, C. C. *Nature* **233**, 240–243 (1981).
- Lorius, C., Merlivat, L., Jouzel, J. & Pourchet, M. *Nature* **280**, 644–648 (1978).
- Barkov, N. I., Gordienko, F. G., Korotkevich, E. S. & Kotlyakov, V. M. *Int. antarct. Exped.* **90**, 39–49 (1975).
- Clapperton, C. H., Sugden, D. E., Birnie, R. V., Hansom, J. D. & Thon *Glacial History and World Palaeoenvironments* (ed. van Zinderen Bakker, 1986).

A film of organic matter at the fresh-water/sea-water interface of an estuary

Vera Žutić & Tarzan Legović*

Center for Marine Research, Institute Rudjer Bošković,
PO 1016, YU-41001 Zagreb, Yugoslavia

In estuaries one finds three types of interface: air/water, fresh-water/sea-water and water/solid. Physical, chemical and biological processes are relatively well understood at the air/water^{1,2} and solid/water interfaces³, which are common to other aquatic systems; processes and, in particular, organic structures at the fresh-water/sea-water interface are less well known. Here we report measurements of surface-active organic material along the vertical profile of a stratified estuary that point to the existence of a well-defined film, and subsequent visual observations by divers which confirm this discovery. The film is formed at the fresh-water/sea-water interface by accumulation and condensation of plankton-derived organic matter under the influence of salinity and shear gradients. It is composed of dissolved and insoluble, liquid, surface-active material. The film contributes to the stability of the fresh-water/sea-water interface, contains a potential food source for heterotrophic organisms and accumulates pollutants. Because we have found the film in an estuary with a low content of organic matter, we expect that it occurs in most salt-wedge estuaries of the world.

Mixing of fresh-water and sea-water induces great changes of ionic strength, chemical composition, speciation of material and distribution of organisms⁴. When river flow is large in areas of low tidal-amplitudes, fresh water flows out of the estuary above a deeper layer of undiluted sea water and this forms a salt wedge⁵.

The estuary of the river Krka on the eastern Adriatic coast of the Mediterranean Sea (Fig. 1) is characterized by low terrigenous input due to a karstic drainage area and a series of travertine barriers that precede the estuary. A high stratification (salinity gradient of 30‰ per 20 cm) reduces the river/sea-water interface to a well-defined, visible region. Water samples at 10 and 25-cm intervals around the interface were collected by scuba divers using 1 l glass or Teflon bottles. The diameter of the orifice (2 cm) determined the vertical resolution of sampling. Here we present a typical set of measurements, taken at a station in the middle of the estuary during the spring cruise of 1985. The first measurements were made in 1982³¹. A complete set of measurements made during 1983–1986 may be found in ref. 6.

Only temperature was measured *in situ*; salinity was measured conductometrically in the samples. Filtration and electrochemical surfactant characterization were completed within 12 h of sampling. Total suspended matter (SM) dried at 40 °C; particulate organic carbon (POC), chlorophyll *a* and phaeophytin were measured after filtration on Whatman GF/F glass-fibre filters. Surface-active organic matter in the samples was characterized as either dissolved or dispersed ('surface-active aggregates') by measuring adsorption at the mercury electrode/sea-water interface^{7–11}. The dropping mercury electrode served as a probe and a model hydrophobic surface with controllable surface charge. Surfactant activity of the samples is expressed in equivalents of synthetic surfactant Triton-X-100 in milligrams per litre^{7–9} and was measured both directly and in filtered samples. The electrochemical response in unfiltered samples allowed direct counting and characterization of fluid surface-active aggregates^{10–11} otherwise retained on the filter. The aggregates rapidly coalesce and reorganize into the adsorbed layer at the liquid interface; they resemble dispersed methyl oleate and are seen as oily droplets of 0.5–10 µm.

As revealed from the salinity profile (Fig. 2a), nearly homogeneous upper, fresh-water, and lower, sea-water, layers are separated by a strong halocline. In the narrow zone around the halocline there is a sharp, and sometimes multiple, discontinuity in most, particularly surface-active properties.

In comparison with other estuaries^{4,12}, concentration of SM and POC in the Krka estuary is low. At this station the net sedimentation rate is 0.016 g cm⁻² yr⁻¹ (ref. 13). Both SM and POC have maximum values at the interface. The content of SM is 80% higher and the content of POC is four times higher (Fig. 2b) than in the rest of the water column.

It is important to note that the chlorophyll *a* concentration in the interface sample (Fig. 2c) is lower than in the adjacent upper fresh-water layer, whereas the concentration of phaeophytin, a chlorophyll degradation product, is 30 times higher. A phytoplankton analysis of the sample showed presence of dead cells and fragments of fresh-water microplankton in addition to heterotrophic nanoplankton (6×10^6 cells l⁻¹) (D. Viličić, V. Ž. and T. L., in preparation).

When compared with other estuaries studied so far^{14–16}, surfactant activity in the fresh-water layer is also low. Surface-active organic matter is predominantly in soluble form with few surface-active aggregates (Fig. 2d). In the sea-water layer below the halocline, surfactant activity is slightly higher than in the upper, fresh-water layer. The electrochemical signal reveals predominantly soluble surfactants and even fewer aggregates than in the fresh-water layer.

In the sample taken at the visible interface, total surfactant activity is increased by a factor of three (Fig. 2d) with respect to the fresh-water or sea-water bulk layers.

Adjacent to the interface one typically finds minima in the concentration of the aggregates, total surfactant activity and POC, indicating a depletion due to transformation and accumulation into the interfacial layer.

Measurement of total mercury concentration showed a maximum similar to that of organic aggregates, whereas maxima for electroactive (dissolved) lead, cadmium, copper and zinc were less pronounced¹⁷.

Vertical distributions, as represented in Fig. 2a–d, are typically obtained by analysing fresh samples taken along the salt wedge under variable meteorological conditions, biological activity, fresh-water flow and time within the tidal cycle⁶. The same phenomenon was also observed in other salt-wedge estuaries in the Mediterranean^{11,16}. By analogy to the air/sea-water interface², the concentration measured in the fresh-water/sea-water interface sample may be expressed as a surface excess (Table 1). The surface-active material present in the interface is

Table 1 Surface excess of various parameters of the organic matter in fresh-water/sea-water interface and a number of equivalent monolayers

Parameter	Surface excess (mg m ⁻²)	Equivalent monolayers
Surfactant activity	20	25
POC	10	12
Phaeophytin	0.56	0.5
Surface-active aggregates	6×10^8	2

sufficient to form several condensed monolayers. It is accumulated by differential settling at the density and the shear gradient where the residence time of solutes and particles is longer than in the rest of the water column. The insoluble surface-active material is formed by condensation of excretion and decomposition products of phytoplankton^{8,10,11}. The process is salting out and supramolecular organization by hydrogen-bonding and hydrophobic attraction, which is known to take place among surfactant molecules, micelles and non-living aggregates, as well as plant cell membranes and organelles when ionic strength in

* Present address: Institute of Ecology, University of Georgia, Athens, Georgia 30602, USA

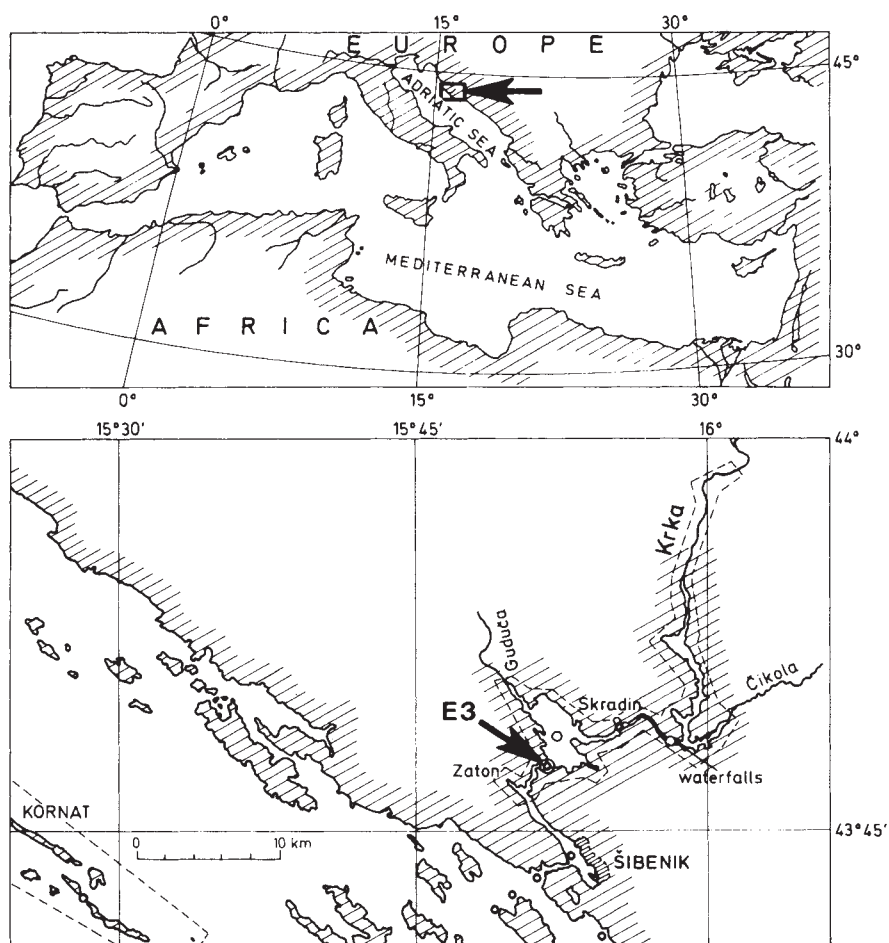
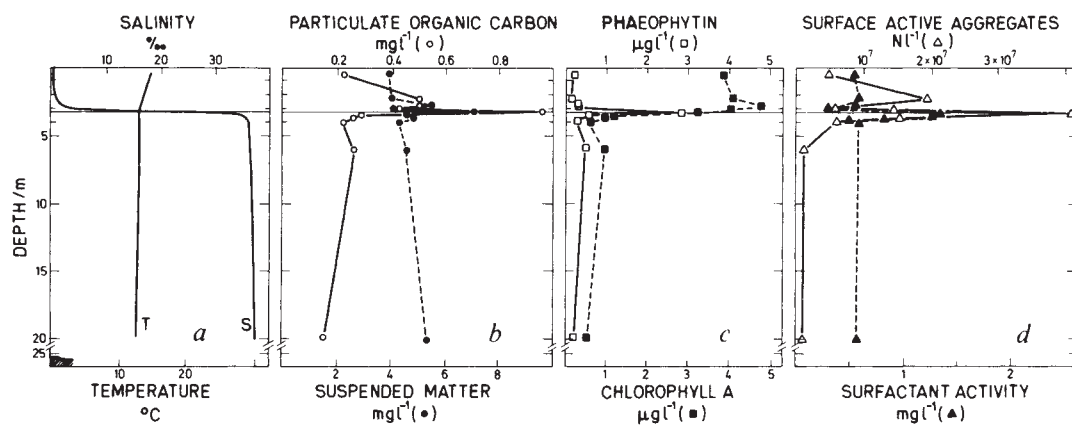


Fig. 1 Krka estuary and measurement station (E3). The coordinates are $43^{\circ}47'30''$ N and $15^{\circ}51'36''$ E (dashed lines indicate boundaries of national parks).

Fig. 2 *a*, Typical vertical profile for salinity and temperature; *b*, particulate organic carbon ($\circ$) and suspended matter ($\bullet$); *c*, phaeophytin ($\square$) and chlorophyll *a* ($\blacksquare$); *d*, surface-active aggregates ($\triangle$) and surfactant activity ($\blacktriangle$). Data from station E3 cca 9-km downstream of the front (beginning) of the salt wedge on 4 April 1985.



an aqueous medium is above 0.1 M^{18-20} . The electrochemical characteristics of this material closely resemble chlorophyll-derived insoluble surface-active complex (glucopeptide-acyl lipid) identified recently at other natural interfaces such as microbubbles^{21,22} and sea-surface microlayers²³.

Laboratory culture experiments showed that bacteria do not play a significant role in production of insoluble liquid surfactants; alone, bacteria did not produce them in substantial amounts⁸, and the production in axenic cultures and cultures of microflagellates with bacteria did not show significant differences^{8,10}. Although the highest in the water column, concentration of bacteria in the interface is still low ($\leq 10^{10}$ cells l^{-1} ; M. Šolić and N. Krstulović, in preparation).

The material that appears in the water sample as an unstable heterodispersion forms a condensed film at the fresh-water/sea-water interface. The vertical position of the film is at the place of the highest salinity gradient. More soluble surface-active

molecules are attached to the insoluble film and extend deeper into the aqueous phases on both sides.

Under still conditions, the rate of reconstitution of the film from dispersion in the sample is $2 \times 10^{-11} \text{ g cm}^{-2} \text{ s}^{-1}$. This rate may serve as the lower bound, because in the interface the turbulence is greater and consequently the transport of aggregates is faster. The relaxation time of coalescence events and spreading of individual aggregates at the interface is between 10 and 100 ms.

The film was visible when observed from the saline and more transparent layer under a high angle. It appears as a silvery-greyish sheen. Similar appearance was reported in the literature for a bimolecular lipid membrane reconstituted from solubilized membrane lipids in 0.1 M saline aqueous solution²⁴ and suspended on a metal ring.

As expected, the estuarine film does not resemble a tightly stretched membrane but a stationary, slightly wavy, surface

moving downstream. The amplitude of the waves is 1–2 cm and the wavelength ~10 cm. The film is stable and self-sealing. When a Plexiglass base of a cylinder is immersed at a high angle to the film plane, 1–3 faint surfaces parallel to the film, extending up to 2–3 cm above it, are observed.

We injected a lipophilic dye, Sudan III, into the saline layer just below the film and observed that the dye did not cross the film but spread laterally along it, forming a visible, yellowish film. In a period of 10–20 s the film gradually lost its artificial colour.

We expect that the film affects several estuarine processes. (1) It stabilizes the density gradient by diminishing vertical turbulent mixing of fresh water and saline water, and also damps the interfacial waves. (2) It should affect energy- and mass-transport processes across the interface, analogously to the film at the air/water interface. (3) It should affect transformation of particles and solutes as all of the suspended material entering the estuary passes through the interface. Because residence time of particles in the interface is larger than elsewhere in the water column^{25,26} (vertical transport is the smallest), particles and solutes have more time to interact with the organic matter. The interactions such as flocculation, complexation and sorption have half times in minutes^{25,27–29}. All of these processes are fast enough to affect the particles crossing the interface. Mineral particles that sink across the film acquire organic coatings^{16,30} and transport the film material to the deeper layer²⁶ whereas light organic particles are accumulated on the film. (4) Organic matter retained in the interface and on the film represents a potential food source for heterotrophic organisms. The film also acts as a scavenger (trap) of pollutants ranging from mercury to chlorinated hydrocarbons^{6,17}. It might thus represent a site for pollutant incorporation into food chains.

We thank S. Macura, Ž. Kwokal and K. Pregernik for underwater sampling and observations; M. Najdek for pigment analysis; Dr. G. Cauwet for POC analysis; Z. Gržetić for hydrographical data; and Dr. J. Marston for improving the manuscript. This research is supported by the Authority for Science Research of SR Croatia, Yugoslavia and UNEP/IOC (Mediterranean Trust Fund).

Received 18 February; accepted 11 May 1987.

- MacIntyre, F. in *The Sea*, Vol. 5 (ed. Goldberg, E. D.) 245–299 (Wiley, New York, 1974).
- Hunter, K. A. & Liss, P. in *Marine Chemistry* (eds Dursma, K. E. & Dawson, R.) 259–298 (Elsevier, Amsterdam, 1981).
- Stumm, W. & Morgan, J. J. *Aquatic Chemistry* (Wiley, New York, 1981).
- Ketchum, B. H. *Estuaries and Enclosed Seas, Ecosystems of the world* Vol. 23 (Elsevier, Amsterdam, 1983).
- Retray, M. Jr & Mitsuda, E. *Estuar. Coast. mar. Sci.* **2**, 375–394 (1974).
- Žutić, V. *Long-Term Research and Pollution Monitoring of the Krka Estuary and Kornati Archipelago (Adriatic Sea)*, MED POL-Phase II (Center for Marine Research, R. Bošković Institute, Zagreb, Annual reports, 1983–86).
- Zvonarić, T., Žutić, V. & Branica, M. *Thalassia jugosl.* **9**, 65–75 (1973).
- Žutić, V., Cosović, B., Bihari, N., Marčenko, E. & Kršinić, F. *Mar. Chem.* **10**, 505–520 (1981).
- Hunter, K. A. & Liss, P. *Water Res.* **15**, 203–215 (1981).
- Žutić, V., Pleše, T., Tomašić, J. & Legović, T. *Molec. Cryst. Liq. Cryst.* **113**, 131–145 (1984).
- Pleše, T. & Žutić, V. *J. electroanal. Chem.* **175**, 299–312 (1984).
- Kennedy, V. S. (ed.) *The Estuary as a Filter* (Academic, Orlando, 1984).
- Juračić, M. thesis, Univ. of Zagreb (1987).
- Hunter, K. A. & Liss, P. *Limnol. Oceanogr.* **25**, 807–822 (1980).
- Hunter, K. A. *Geochim. cosmochim. Acta* **47**, 467–473 (1983).
- Žutić, V., Cauwet, G. & Monaco, A. in *Proc. 7th Workshop on Marine Pollution of the Mediterranean, Luzern, 1984*, 173–180 (UNEP, 1986).
- Martinčić, D., Branica, M. & Kwokal, Ž. in *Long-Term Research and Pollution Monitoring of the Krka Estuary and Kornati Archipelago (Adriatic Sea)*, MED POL-Phase II (Center for Marine Research, R. Bošković Institute, Zagreb, Annual reports, 1983–86).
- Tanford, C. *The Hydrophobic Effect* (Wiley, New York, 1980).
- Istraelachvili, J. N. & Pasey, R. M. *J. Colloid Interface Sci.* **98**, 500–514 (1984).
- Staehelin, L. A. & Arntzen, C. J. in *Chloroplast Organization and Energy Transfer in Photosynthesis* (ed. Porter, G.) 147–175 (Ciba Foundation Symposium No. 61, Amsterdam, 1978).
- D'Arrigo, J. S., Saiz-Jimenez, C. & Reiner, N. S. *J. Colloid Interface Sci.* **100**, 96–105 (1984).
- D'Arrigo, J. S. *J. Colloid Interface Sci.* **100**, 106–111 (1984).
- Berger, W. R. & Means, J. C. in *Marine and Estuarine Geochemistry* (eds Sigelo, A. C. & Hattori, A.) 47–67 (Lewis, Chelsea, 1985).
- Mueller, P., Rudin, D. O., Tien, T. H. & Wescott, W. C. *Nature* **194**, 979–980 (1962).
- Sakamoto, W. *Bull. Ocean. Res., Inst. Univ. Tokyo*, **5**, 1–49 (1972).
- Aloisi, J. C. et al. *Oceanol. Acta* **5**, 481–491 (1982).
- Mayer, L. M. *Geochim. cosmochim. Acta* **46**, 1527–1536 (1982).
- Raspor, B., Nuernberg, N. W., Valenta, P. & Branica, M. *Limnol. Oceanogr.* **26**, 54–66 (1981).
- Harvey, R. W., Lion, L. W., Young, L. Y. & Leckie, J. O. *J. mar. Res.* **40**, 1201–1211 (1982).
- Žutić, V. & Tomašić, J. *Mar. Chem.* (in the press).
- Cosović, B., Žutić, V. & Cauwet, G. *28th CIESM Meeting, Cannes* (November, 1983).

Localization of the gene for familial adenomatous polyposis on chromosome 5

W. F. Bodmer*, C. J. Bailey*, J. Bodmer†, H. J. R. Bussey‡, A. Ellis§, P. Gorman||, F. C. Lucibello*, V. A. Murday¶, S. H. Rider||, P. Scambler#, D. Sheer||, E. Solomon** & N. K. Spurr††

* Director's Laboratory, † Tissue Antigen, ** Somatic Cell Genetics and || Human Cytogenetics Laboratories, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

‡ Polyposis Registry, ¶ ICRF Colorectal Unit, St Mark's Hospital, City Road, London EC1V 2PS, UK

§ Gastroenterology Unit, Broadgreen Hospital, Liverpool L14 3LB, UK

Department of Biochemistry, St Mary's Hospital Medical School, London W2 1PG, UK

†† Human Genetic Resources Laboratory, ICRF Clare Hall Laboratories, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3LD, UK

Colorectal cancer is the second most common cancer in the United Kingdom and other developed countries in the West. Although it is usually not familial, there is a rare dominantly inherited susceptibility to colon cancer, familial adenomatous polyposis (FAP; also often previously called familial polyposis coli). During adolescence affected individuals develop from a few hundred to over a thousand adenomatous polyps in their large bowel. These are sufficiently likely to give rise to adenocarcinomas to make prophylactic removal of the colon usual in diagnosed FAP individuals. Adenomas may occur elsewhere in the gastrointestinal tract and the condition is often associated with other extracolonic lesions, such as epidermoid cysts, jaw osteomata and fibrous desmoid tumours^{1–4}. Adenomata have been suggested to be precancerous states for most colorectal tumours^{5,6}. Knudson⁷ has suggested that the mutation for a dominantly inherited cancer susceptibility may be the first step in a recessive change in the tumour cells, and that the same gene may be involved in both familial and non-familial cases of a given tumour. Following up a case report of an interstitial deletion of chromosome 5 in a mentally retarded individual with multiple developmental abnormalities and FAP⁸, we have now shown that the FAP gene is on chromosome 5, most probably near bands 5q21–q22.

Most of the families for the study were identified by the Polyposis Register at St Mark's Hospital, London and by the Gastroenterology Unit, Broadgreen Hospital, Liverpool, and were well characterized with respect to clinical, pathological and pedigree information. Sterile blood samples were collected from these families and lymphocytes separated out and frozen. These were subsequently transformed by Epstein-Barr virus (EBV) to produce a permanent source of DNA (ref. 9). The sediment of these mononuclear cell separations was also used to provide an immediate source of DNA (Fig. 1 legend). Five polymorphic DNA probes (see legends to Figs 1 and 2) assigned to chromosome 5 and mostly not regionally localized, were tested on up to 124 members from 13 different families. One of

Table 1 Lod scores from two-point analyses for linkage between C11p11, L1.4 and FAP

Marker loci	Recombination fraction (θ)				
	0.0	0.10	0.20	0.30	0.40
C11p11-FAP	3.26	2.56	1.88	1.15	0.44
L1.4-FAP	−2.50	0.78	0.66	0.37	0.12

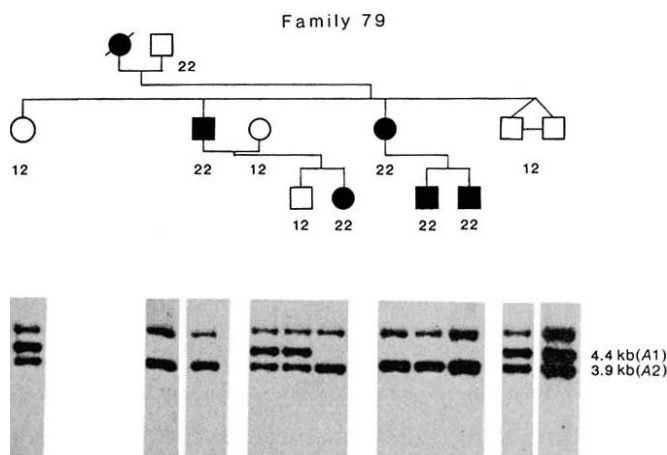


Fig. 1 Family 79 demonstrating segregation of FAP with a 3.9-kb fragment using the probe C11p11, a 3.6-kb *EcoRI* fragment cloned into pUC8 that reveals a *TaqI* polymorphism producing alleles A1 at 4.4 kb and A2 at 3.9 kb. Symbols: □, male; ○, female; solid symbol, affected individuals.

Methods. Blood samples were collected from family members and lymphocytes separated and frozen for future transformation with EBV (ref. 9). DNA of high relative molecular mass was extracted from both the residue of the separation and from the transformed lymphocytes. The cells were lysed in a solution containing 0.33 M sucrose, 5 mM MgCl₂, 10 mM Tris-HCl pH 7.5, 1% Triton X-100. Nuclei were pelleted at 8,000g at 4 °C for 10 min and resuspended in 75 mM NaCl, 25 mM EDTA, 200 µg ml⁻¹ proteinase K and 0.5% SDS. The mixture was incubated overnight at 37 °C. The aqueous phase was extracted three times in phenol buffered with 2 mM Tris and 2 mM EDTA and once in isoamyl alcohol:chloroform (1:24). The DNA was reprecipitated with 1/10 vol. 3 M sodium acetate pH 5.2 and two vols absolute alcohol and then resuspended in TE. Aliquots (10 µg) of genomic DNA were digested overnight with 50 units of restriction enzyme, under the recommended conditions (*TaqI* digests for 4–6 h). The samples were run in 0.8% (w/v) agarose gels for 36–48 h, at 1–1.5 V cm⁻¹. The gels were depurinated in 0.25 M HCl for 10 min, denatured in NaOH for 45 min and then neutralized for 60 min. The DNA was then transferred to Hybond-N nylon filters (Amersham), following the manufacturer's recommendation. Probes were labelled with [α -³²P]dCTP by the oligolabelling method¹⁷ for 4–6 h at 37 °C. The filters were prehybridized in 6×SSC, 0.1% SDS, 5× Denhardt's solution containing 25 µg ml⁻¹ of denatured salmon sperm DNA for 1–2 h at 65 °C. Hybridization was carried out overnight at 65 °C in the same solution but supplemented with 10 µCi ml⁻¹ of labelled probe and 10% dextran sulphate.

the probes, C11p11 (provided by the Department of Biochemistry, St Mary's Hospital, London), showed evidence for close linkage to FAP. This probe reveals a *TaqI* polymorphism, producing alleles of 3.9 kilobases (kb) and 4.4 kb with frequencies of about 0.92 and 0.08 (based on 53 individuals, see Fig. 1). Six of the families, including 79 typed individuals, were informative and these gave a combined maximum lod score of 3.26 at a recombination fraction of $\theta = 0$, using the LIPED linkage analysis computer program¹⁰. The lod scores for the linkage of FAP to L1.4 (*D5S4*) (ref. 11) are also given and, although not significant, give an estimate for the recombination fraction θ of about 0.1 (see Table 1). The other probes (L1.7 (*D5S1*) (ref. 11) from P. Pearson; λ MS8 from A. Jeffreys (ref. 12) and pHexXbaI, a hexosaminidase B probe from R. Gravel) showed no significant linkage with FAP or with each other. There was no obvious heterogeneity between the families with respect to the linkage of C11p11 with FAP. Assigning the most probable genotypes in all cases and assuming close linkage, the data are consistent with no recombinants amongst 24 meioses, with no genotype predictions being necessary in 12 out of these 24 meioses.

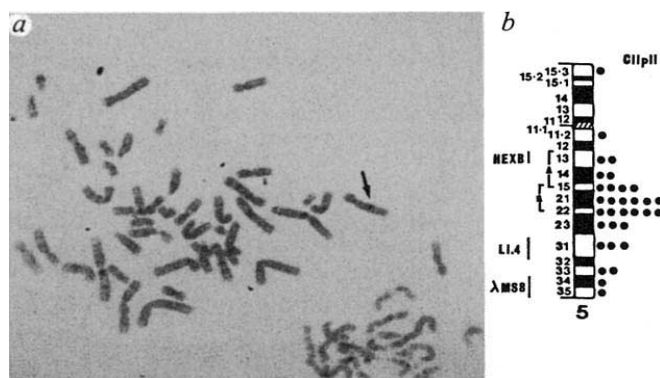


Fig. 2 *a*, Representative chromosome spread showing *in situ* hybridization of probe C11p11. *In situ* hybridization was performed on chromosomes prepared from PHA-stimulated normal blood lymphocytes essentially as described by Harper and Saunders¹⁸. Probe C11p11 was labelled by random oligonucleotide priming and ³H-deoxyribonucleotides to a specific activity of 1×10^7 c.p.m. µg⁻¹. Slides were washed in 2×SSC at 39 °C, dehydrated and dipped in Ilford K-5 emulsion. They were developed after 10 days. *b*, Fine chromosome mapping of probe C11p11 and putative regional assignment of probes L1.4 and λ MS8. Fine chromosome mapping was carried out with the C11p11 probe; 221 silver grains were scored on chromosomes from 150 metaphases, with 22 grains (9.9%) present over bands 5q15–q23 and most over bands 5q21–22. Preliminary data for *in situ* hybridization of the L1.4 probe localizes it to band 5q31. Localization of λ MS8 was by *in situ* hybridization (N. J. Royle and A. J. Jeffreys, personal communication). The symbols ▲ and the associated regions represent the two alternative positions for the interstitial deletion reported by Herrera *et al.*⁸ L1.7 (*D5S1*, ref. 10) is a 1.14-kb *EcoRI* fragment cloned into pBR322 and reveals a *BglII* polymorphism with alleles A1 at 7.9 kb and A2 at 10.6 kb. L1.7 also reveals a *PstI* polymorphism with alleles B1 at 2.8 kb and B2 at 2.3 kb. L1.4 (*D5S4*, ref. 10) is a 0.72-kb *EcoRI* fragment cloned into pBR322 and reveals an *EcoRI* polymorphism with alleles A1 at 0.7 kb and A2 at 0.6 kb. L1.4 also reveals an *RsaI* polymorphism with alleles C1 at 2.4 kb and C2 at 2.2 kb (ref. 19). Probe λ MS8 (ref. 11) produces multiple alleles with *HinfI*. Probe pHexXbaI (ref. 10) is a 1.5-kb *XbaI* cDNA fragment from the hexosaminidase B gene, cloned into pSP64. A 1.1-kb *BamHI*-*PstI* fragment reveals a *PstI* polymorphism with alleles A1 at 1.6 kb and A2 at 1.0 kb.

To relate the linkage data to the interstitial deletion reported by Herrera *et al.*⁸, *in situ* hybridization was carried out to localize C11p11 and L1.4 on chromosome 5. Probe C11p11 localizes to 5q21–q22 (Fig. 2a) which is consistent with the deletion described by Herrera *et al.* being at 5q15–q22 rather than their cytogenetically indistinguishable alternative 5q13–q15. This localization is also consistent with a case report of a large bowel adenocarcinoma which has a 5q12?–q22 deletion in addition to other cytogenetic abnormalities¹³. The lack of significant linkage between *HEXB* or λ MS8 and FAP is consistent with their having been localized on the long arm of chromosome 5 (Fig. 2b). In the case of L1.4 this may be due to the small number of informative families rather than its location.

The demonstration in the accompanying paper¹⁴ that >20% of sporadic colorectal carcinomas become homo- or hemizygous for chromosome 5 alleles, the appreciable lod score for linkage with C11p11 and the coincidental localization of the probe with the two deletions^{8,13} provides overwhelming evidence for assigning FAP to chromosome 5, probably to 5q21–q22.

The marker loss data, together with the localization of FAP, extend Knudson's ideas to FAP and, at least, to a major subset

of colorectal carcinomas. The data thus suggest that becoming recessive for a gene around 5q21-q22 is a key step in the progression of many colorectal carcinomas, both sporadic and familial. The fact that polyps from FAP patients do not yet show the recessive change, as shown in the accompanying paper¹⁴, is consistent with data using G6PD markers which suggested that polyps were not clonal¹⁵.

One possible explanation for how heterozygosity for a deficiency can give rise to localized growth abnormalities such as polyps, is through a threshold effect involving, for example, negative control over the production of growth factors. In this model, the normal homozygote produces enough of the FAP gene product so that random fluctuations of its concentration never, or at least only very rarely, go below a threshold that permits localized excessive growth. The deficient heterozygote, however, may allow random fluctuations in the level of the FAP gene product to go below this threshold relatively frequently. In those areas of the colon where this happens, excessive epithelial growth may be initiated, giving rise to polyps. Once initiated, this growth may persist by, for example, some form of feedback control for production of a growth factor. This is at least one way to explain how a gene dosage effect can give rise to discontinuous growths, namely the polyps, which look as though they should be clonal. Presumably the polyps, once they have arisen, provide the opportunity for the second, recessive, change to take place, followed no doubt by other changes which lead to the overt carcinoma.

The apparent similarity in frequency of FAP throughout the world, together with the relatively high proportion of cases (up to 40%) which may be sporadic, due to new mutations¹, suggests that most FAP families may carry different mutations. At this detailed genetic level the disease may therefore be quite heterogeneous. The incidence of FAP is estimated to be at least 1/10,000 and perhaps up to 1/5,000, which suggests that the mutation rate to FAP is at least 1/100,000 and could be as high as 1/25,000. Perhaps there are sequences around the FAP gene which predispose it to relatively high mutation rates.

Now that the FAP gene has been localized, long-range DNA analysis using pulsed field gel electrophoresis and chromosome jumping techniques¹⁶, together with searches for appropriate epithelial-specific messages encoded in this genetic region, should lead to the identification of the FAP gene and its function. This will not only provide a basis for the presymptomatic or prenatal diagnosis of FAP, but may also lead to approaches to counter-acting the FAP defect and provide new clues to specific treatments for at least some forms of sporadic colon carcinoma.

The following centres and individuals collaborated in this study by providing family samples and probes; we thank them for their cooperation and hard work. Steven Bryant, Human Genetics Resources Laboratory, ICRF Clare Hall Laboratories; Dr R. Cartwright, Mrs J. Greatrex, Mrs P. Roberts, Dr L. Williams, the Leukaemia Research Fund Centre for Clinical Epidemiology, Leeds University, Leeds; Dr R. Gravel, Hospital for Sick Children, Toronto, Canada; Dr A. Jeffreys, Department of Genetics, University of Leicester; Lorna Kennedy and Steve Marsh, Tissue Antigen Laboratory, Imperial Cancer Research Fund, London; Paul A. Lawson, Director's Laboratory, ICRF, London; Dr C. McKeown and Professor R. H. Harris, Department of Medical Genetics, St Mary's Hospital, Manchester; Dr P. L. Pearson, Department of Human Genetics, State University, Leiden, The Netherlands; Dr B. Ponder, Institute of Cancer Research, Royal Marsden Hospital, Sutton; Dr S. Ritchie and Miss K. Neale, Polyposis Registry, St Mark's Hospital, London; Sister J. Shaw, Gastroenterology Unit, Broadgreen Hospital, Liverpool.

Received 22 June; accepted 21 July 1987.

1. Bulow, S. *Dan. med. Bull.* **34**, 1-15 (1987).
2. Veale, A. M. O. *Intestinal Polyposis* (Cambridge University Press, 1965).
3. Bussey, H. J. R. *Familial Polyposis Coli* (Johns Hopkins University Press, 1975).

4. Gardner, E. *Am. J. hum. Genet.* **3**, 167-176 (1951).
5. Muto, T., Bussey, H. J. R. & Morson, B. C. *Cancer*, **36**, 2251-2270 (1975).
6. Morson, B. C. *Clin. Radiol.* **35**, 425-431 (1984).
7. Knudson, A. G. Jr *Proc. natn. Acad. Sci. U.S.A.* **68**, 820-823 (1971).
8. Herrera, L., Kakati, S., Gibas, L., Pietrzak, E. & Sandberg, A. A. *Am. J. med. Genet.* **25**, 473-476 (1986).
9. Miller, G. & Lipman, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 190-194 (1973).
10. Ott, J. *Am. J. hum. Genet.* **20**, 588-597 (1974).
11. 8th Int. Workshop on Human Gene Mapping *Cytogenet. Cell Genet.* **40**, 1-824 (1985).
12. Wong, Z. *et al. Ann. hum. Genet.* (in the press).
13. Ferti-Passantonopolou, A., Panani, A., Avgerinos, A. & Raptis, S. *Cancer Genet. Cytogenet.* **21**, 361-364 (1986).
14. Solomon, E. *et al. Nature* **328**, 616-619 (1987).
15. Hsu, S. H., Luk, G. D., Krush, A. J., Hamilton, S. R. & Hoover, H. H. Jr *Science*, **221**, 951-953 (1983).
16. Poustka, A., Pohl, T. M., Barlow, D. P., Frischauf, A. & Lehrach, H. *Nature* **325**, 353-355 (1987).
17. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
18. Harper, M. E. & Saunders, G. F. *Chromosoma* **83**, 431-439 (1981).
19. Bailey, C. J., Rider, S. H. & Bodmer, W. F. *Nucleic Acids Res.* (in the press).

Chromosome 5 allele loss in human colorectal carcinomas

E. Solomon*, R. Voss†, V. Hall*, W. F. Bodmer‡, J. R. Jass§, A. J. Jeffreys||, F. C. Lucibello‡, I. Patel|| & S. H. Rider¶

* Somatic Cell Genetics Laboratory, † Director's Laboratory, ¶ Human Cytogenetics Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

‡ Department of Human Genetics, Hadassah Medical School, Jerusalem, Israel

§ St Mark's Hospital, City Road, London EC1V 2PS, UK

|| Department of Genetics, University of Leicester, Leicester LE1 7RH, UK

That the sporadic and inherited forms of a particular cancer could both result from mutations in the same gene was first proposed by Knudson¹. He further proposed that these mutations act recessively at the cellular level, and that both copies of the gene must be lost for the cancer to develop². In sporadic cases both events occur somatically whereas in dominant familial cases susceptibility is inherited through a germline mutation and the cancer develops after a somatic change in the homologous allele. This model has since been substantiated in the case of retinoblastoma³, Wilms' tumour⁴⁻⁷, acoustic neuroma⁸ and several other tumours^{9,10}, in which loss of heterozygosity was shown in tumour material compared to normal tissue from the same patient. The dominantly inherited disorder, familial adenomatous polyposis (FAP, also called familial polyposis coli), which gives rise to multiple adenomatous polyps in the colon that have a relatively high probability of progressing to a malignant adenocarcinoma¹¹, provides a basis for studying recessive genes in the far more common colorectal carcinomas using this approach. Following a clue as to the location of the FAP gene given by a case report of an individual with an interstitial deletion of chromosome 5q, who had FAP and multiple developmental abnormalities¹², we have examined sporadic colorectal adenocarcinomas for loss of alleles on chromosome 5. Using a highly polymorphic 'minisatellite' probe¹³ which maps to chromosome 5q we have shown that at least 20% of this highly heterogeneous set of tumours lose one of the alleles present in matched normal tissue. This parallels the assignment of the FAP gene to chromosome 5 (see accompanying paper¹⁴) and suggests that becoming recessive for this gene may be a critical step in the progression of a relatively high proportion of colorectal cancers.

Tumour material and adjacent normal mucosa samples were collected from about 45 cases of colorectal carcinomas, mostly from St Mark's Hospital, London. DNA preparations from these matched normal-tumour pairs were analysed for restriction fragment length polymorphism (RFLP) differences with a variety of probes that had been assigned to particular chromosomes. Following the clue that the FAP gene might be on chromosome 5, attention was concentrated on probe λMS8 (ref. 13) which

Table 1 Allele losses in colorectal carcinomas with chromosome-5 probes λ MS8 and L1.4

Patient	Tumour type	Dukes staging ²¹	Probes			
			λ MS8 <i>HinfI</i> (alleles detected)		L1.4 <i>EcoRI</i> (alleles detected)	
			N	T	N	T
1	AC rectum	C1	2	2	12	12
2	AC sigmoid colon	B	2	2	12	—
3	AC rectum	C1	2	2	—	—
4	AC rectum	C1	2	(1)	—	—
5	AC rectum	C1	—	—	12	2
6	AC rectum	B	2	2	12	12
7	AC rectum	C2	2	2	—	—
8	AC rectum	B	2	2	—	—
9	AC sigmoid colon	B	2	1	12	(2)
10	AC rectum	A	2	2	—	—
11	AC rectum	A	2	(1)	12	(1)
12	AC sigmoid colon	C1	2	2	—	—
13	AC rectum	C1	2	2	12	—
14	AC rectum	B	2	2	—	—
15	AC asc. colon	B	2	2	12	12
16	AC rectum	C1	2	2	—	—
17	Mucinous AC rectum	C1	2	1	—	—
18	AC rectum	B	2	1	—	—
19	AC rectum	B	2	1	12	2
20	AC rectum	C1	2	2	—	—
21	AC sigmoid colon	B	2	(1)	12	12
22	Mucinous AC asc. colon	B	2	2	12	12
23	AC rectum	B	2	2	—	—
24	AC rectum	B	2	2	—	—
25	AC sigmoid colon	B	2	2	—	—
26	AC rectum	C1	2	2	—	—
27	AC rectum	C1	2	2	—	—
28	AC rectum	A	2	2	12	2
29	AC asc. colon	C1	2	2	—	—
30	AC rectum	C1	2	2	—	—
31	Mucinous AC asc. colon	B	2	2	—	—
32	AC rectum	B	2	2	—	—
33	AC rectum	C1	2	2	—	—
34	AC sigmoid colon	B	2	2	—	—
35	AC sigmoid colon	B	2	1	12	2
36	AC rectum	B	2	2	—	—
37	AC sigmoid colon	B	2	1	—	—
38	AC sigmoid colon	B	2	1	—	—
39	AC rectum	C1	2	2	12	12
40	AC sigmoid colon	C1	2	2	12	12
41	AC rectum	B	2	2	—	—
42	AC desc. colon	B	2	2	12	12
43	AC desc. colon	C	2	2	12	12
44	AC rectum	B	—	—	12	12
45	AC asc. colon	B	2	2	12	12

Summary of cases where both probes are informative

λ MS8	L1.4	λ MS8	L1.4	λ MS8	L1.4	λ MS8	L1.4	Total
no loss	no loss	no loss	no loss	no loss	no loss	no loss	no loss	
9	1	1	4	15				

Abbreviations: AC, adenocarcinoma; asc., ascending; desc., descending. Tumour material and adjacent normal mucosa from colorectal cancer patients, obtained during surgery, was frozen immediately in liquid nitrogen. DNA was analysed by standard methods (see Fig. 1 legend). Probe λ MS8 detects alleles of different sizes in almost every individual so that each sample is indicated as having two alleles or only one rather than particular allele sizes. The L1.4 alleles are 0.7 kilobases (kb) (allele 1) and 0.6 kb (allele 2). Symbols: —, non-informative sample, due to lack of heterozygosity; blank space, no information. Tumour samples with less certain allele loss are shown with the most prominent remaining allele in parentheses.

is a locus-specific minisatellite probe that maps to chromosome 5q34-qter (N. J. Royle and A.J.J., unpublished). About 90% of individuals (55 out of 60 in our data) are heterozygous with this probe, showing two single band alleles with a *HinfI* digest. This high level of heterozygosity makes this probe particularly valuable for this type of study. Forty-four heterozygous normal-tumour pairs were examined with this probe using *HinfI* digests (Table 1). Of these, seven tumours show clear allele loss (nos 9, 17, 18, 19, 35, 37 and 38; for representative examples see Fig. 1a) and three others showed obvious but less extreme changes

Table 2 Heterozygosity in colorectal carcinomas with chromosome-specific probes

Chromosome	Probe	Locus	No. of patients tested	No. of heterozygotes		Ref.
				Normal	Tumour	
7	pJ3.11 (<i>MspI</i>)	<i>D7S8</i>	17	9	9	22
	pAg3 (<i>HinfI</i>)	<i>D7S22</i>	9	9	9	13
8	λ 287 (<i>TaqI</i>)	<i>D8S8</i>	33	14	14	23
11	λ 32-1 (<i>MspI</i>)	<i>D11S16</i>	39	19	19	24
12	pPstI (<i>HindIII</i>)	<i>COL2A1</i>	31	17	16	25
	p12-16 (<i>EcoRI</i>)	<i>D12S2</i>	27	4	4	26
13	p7F12 (<i>MspI</i>)	<i>D13S1</i>	26	13	13	27
	p9D11 (<i>MspI</i>)	<i>D13S2</i>	44	24	24	27
	p1E8 (<i>MspI</i>)	<i>D13S4</i>	37	22	22	27
17	c-erbA1 (<i>PvuII</i>)	<i>ERBA1</i>	20	1	1	28
	pHF12-1 (<i>MspI</i>)	<i>D17S1</i>	41	15	15	29

Tumour and normal DNA from colorectal cancer patients, hybridized with probes on chromosomes 7, 8, 11, 12, 13 and 17. Methods are as in Fig. 1. DNA was digested with the enzymes indicated in parentheses next to the probe name. Patients showing loss of chromosome-5 alleles (Table 1) were informative, but showed no loss for other probes as follows: patient 4, *D7S8*, *D13S2*; patient 5, *D13S1*, *D17S1*; patient 9, *D11S16*, *D12S2*, *D17S1*; patient 11, *D13S4*; patient 17, *D7S22*; patient 18, *D13S2*, *D17S1*; patient 19, *D13S2*, *D13S4*; patient 21, *D12S2*, *D13S2*; patient 28, *D11S16*; *D13S2*, *D13S4*, *D17S1*.

in the ratio of one allele to the other when compared with normal tissue (nos 4, 11 and 21; for example see no. 4, Fig. 1a). Allele loss in these latter three cases is presumably partially obscured by relatively large amounts of contaminating normal tissue in the tumour sample. Overall, 23% of tumours showed partial or complete loss of an allele at this locus.

The same normal-tumour pairs were examined with a second probe, L1.4 which maps approximately to 5q31 (refs 14 and 15). The RFLP allele frequencies detected with this probe, using *EcoRI* are, in our data, 0.27 and 0.73 based on 43 individuals. Seventeen informative pairs were studied and, of these, four showed clear allele loss (nos 5, 19, 28 and 35; Fig. 1a) whereas two others showed an obvious change in the ratio of one allele to the other (nos 9 and 11). Again, ambiguous results are probably due to contaminating normal tissue in the tumour sample. The overall results with this probe, namely 6 presumed losses out of 17, were similar to those with λ MS8. There are not enough data to analyse any possible relationship between Dukes stage, or tumour type or location and allele loss.

The results for cases in which both markers λ MS8 and L1.4 were informative are summarized at the bottom of Table 1. A total of 6 out of 15, or 40%, showed changes for one or other marker, although with such small numbers, this is not significantly higher than the 10 out of 44 for λ MS8 alone. Four out of six were changed for both markers, suggesting a high frequency of overall chromosomal loss. A more localized deletion or somatic recombination would explain no. 28 with loss at L1.4 but not λ MS8. One case, no. 21, with loss at λ MS8 but not L1.4, is inconsistent with these interpretations given the presumed map order, centromere-*FAP*-L1.4- λ MS8 (ref. 14). This could, once again, be the result of misclassification for one of the markers due to contaminating normal tissue in the tumour material.

To obtain a more objective quantitative assessment of the allele ratios in the material given in Table 1, the autoradiographs for λ MS8 were scanned with a densitometer. From the scans, the ratios of the amounts of the two alleles in tumour material to those in the matched tissues could be estimated. For true allele loss, this derived ratio should be 0, but experimental error and normal tissue contaminating the tumour will lead to a distribution which should be bimodal. Preliminary results support this prediction and indicate nearly 40% (15 out of 39 cases scanned) have tumour/normal ratios in the lower part of this distribution (Fig. 2). This suggests that a more careful assessment of allele loss might lead to a higher estimate of the proportion

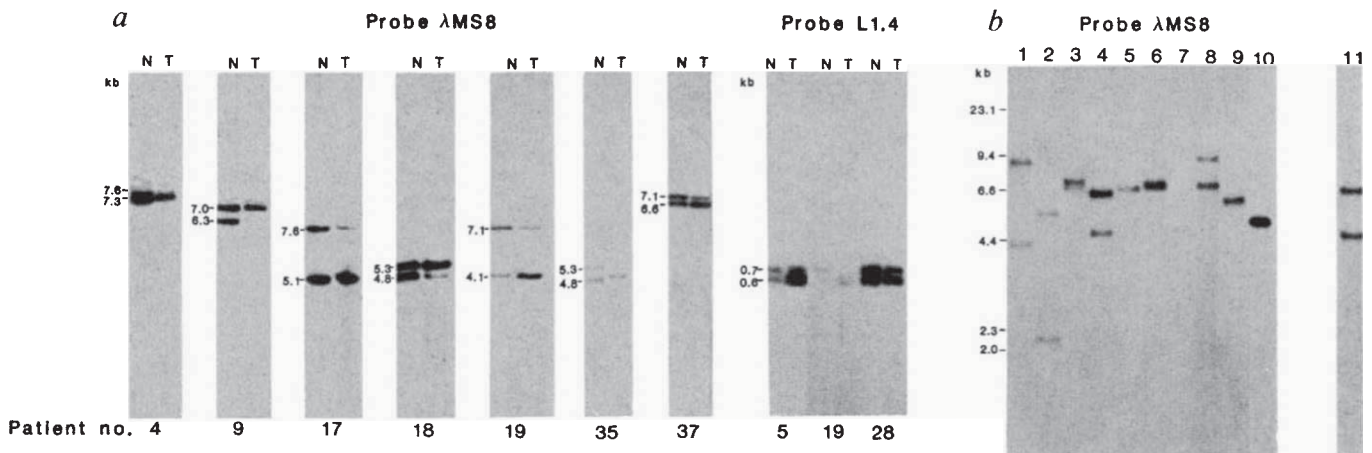


Fig. 1 Allele loss on chromosomes in primary colorectal tumours and cell lines. *a*, Primary colorectal tumours. DNA from matched normal (N)-tumour (T) pairs hybridized with λ MS8 and L1.4. Patient numbers refer to Table 1. *b*, DNA from colorectal cell lines hybridized with λ MS8. Cell lines are: 1, Coll 2/23; 2, Coll 5/3; 3, HT29; 4, LoVo; 5, LS174T; 6, SW48; 7, SW620; 8, SW837; 9, SW1222; 10, PC/JW; 11, PC/AA. References for cell lines as follows: lines 1 and 2, see ref. 30; line 3, ref. 31; line 4, ref. 32; line 5, ref. 33; lines 6-9, ref. 34, lines 10 and 11, ref. 19.

Methods. DNA was extracted from tumour biopsies (T) and normal adjacent mucosa (N) by standard methods³⁵. For analysis with λ MS8, DNA was digested with *Hinf*I (New England Biolabs) and fractionated by agarose gel electrophoresis on 0.8% gels, for 60 h at 20 mA; for L1.4 DNA was digested with *Eco*RI (Boehringer Corp.) and fractionated on 1.2% agarose gels for 18 h at 63 mA. In both cases DNA was transferred to Hybond-N nylon membranes (Amersham) by standard procedures³⁶. Probes were oligo-labelled³⁷ with [α -³²P]dCTP. Hybridization and washing were by standard methods³⁵. Hybridization of λ MS8 was with sodium phosphate and SDS¹³.

of tumours which have lost a chromosome-5 allele. Those tumours showing allele loss which was easily detected by eye on the autoradiographs are indicated by circles in Fig. 2.

To determine whether the loss of alleles on chromosome 5 is specific and not simply a reflection of random loss, polymorphic markers on six other chromosomes, namely 7, 8, 11, 12, 13, 17 were tested. These were chosen because changes in chromosomes 7, 8, 12 and 17 have been reported in colorectal tumours¹⁶⁻¹⁹. These tumours are difficult to karyotype and often show gross rearrangements and aneuploidy with no very specific chromosomal changes. Chromosomes 13 and 11 were tested because of their involvement in retinoblastoma and Wilms' tumours respectively. The results for these other markers are shown in Table 2. In only one case was heterozygosity not maintained, and this involved 1 tumour (out of 21) which showed the loss of an allele on chromosome 12. In addition, ten informative normal-tumour pairs associated with the inherited cancer syndrome MEN2 showed no loss of alleles on chromosome 5 using probe λ MS8 (ref. 20). Thus, the loss of alleles appears to be chromosome-specific and possibly specific to this tumour.

Ten established cell lines derived from sporadic colonic adenocarcinomas and from a polyp from an FAP case were also tested with probe λ MS8. For the colorectal lines, matching normal tissue was not available, but because the predicted frequency of homozygotes is only 1 in 10, any increase beyond this could be significant. The data in Fig. 1*b* show that four of these ten lines show allele loss (nos 5, 7, 9 and 10) and a fifth (no. 6) most probably so, strongly suggesting loss of all or part of chromosome 5 as a significant genetic change in a remarkably high proportion of these tumours. The data on allele loss from cell lines are much clearer than for the fresh tumour material, as might be expected, because for the lines there is no problem of contamination by normal tissue. The χ^2 with continuity correction is 9.0 for the difference between 5/10 and the control λ MS8 homozygote frequency of 5/60 and this is highly significant with $P=0.0035$.

Now that FAP patients are treated by removal of the colon at a relatively early age to prevent cancers developing, it is not easy to obtain colorectal tumour material from FAP patients. We have, therefore, not yet been able to extend our observations to familial cases of colorectal tumours. The cell line PC/AA

(ref. 19), derived from a polyp, did not show loss of a λ MS8 allele. This is consistent with tests of premalignant polyps from 11 FAP patients using λ MS8, which showed no loss of heterozygosity in any case. The recessive change on chromosome 5 is, thus, secondary to the initial development of a polyp (see ref. 14 for further discussion).

The overall data clearly show that a relatively high proportion of sporadic colorectal carcinomas lose all or part of one chromosome 5, presumably because loss of the FAP gene function favours tumour progression. The fact that λ MS8, which is near the tip of 5q, was lost, together with the comparatively common concomitant loss of an allele of the L1.4 locus, suggests that whole chromosome loss by non-disjunction, presumably following an initial mutation on one of the two chromosomes 5, is the

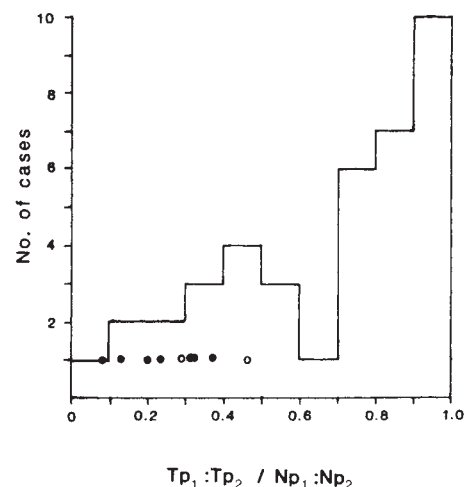


Fig. 2 Densitometric scanning of autoradiographs with normal and tumour DNA hybridized with λ MS8. Autoradiographs were scanned with a Joyce-Loebl Chromoscan 3. Areas under the peaks are plotted as the ratio of this value for the two tumour alleles, Tp_1/Tp_2 , divided by the ratio for the two normal alleles Np_1/Np_2 . ●, The seven cases which could be seen as clear allele loss (left to right: patients 9, 18, 19, 37, 17, 38, 35); ○, those with less obvious loss (patients 21 and 11).

commonest mechanism for becoming recessive, as has been found, for example, for retinoblastoma³. It would be expected that this loss would often be accompanied by reduplication, because monosomy is likely to be disadvantageous for the growth of any cell, and non-disjunction is a relatively common event. Preliminary karyotyping of several cell lines supports this (data not shown). Even if technical developments increase the proportion of tumours in which chromosome 5 allele loss can be demonstrated to around 40%, there will still be a sizeable fraction of tumours in which no change on chromosome 5 is readily detectable. This could in part, of course, be due to relatively localized genetic mechanisms for producing the secondary change, such as point mutations, or small deletions. There is however, the further important likelihood that colorectal tumours are aetiologically heterogeneous with respect to the function of the FAP gene on chromosome 5. Those in which this activity is lost may represent a particular sub-set of colorectal tumours and this may have important implications for diagnosis and treatment. The combined studies of this and the accompanying paper on the localization of the FAP gene and its probable function in a relatively high proportion of colorectal carcinomas, uncovers for the first time an important recessive genetic effect in one of the commonest of all cancers.

We thank the staff of St Mark's Hospital for collecting FAP polyp material, and this staff as well as L. Bobrow and the Histopathology Unit at University College, London for colorectal tumour material. We thank the following for supplying the probes: J. Schmidtke (*D7S8*) (Institut für Humangenetik der Universität, Göttingen, Germany); N. Spurr and J. Brennan (*D8S8*) (ICRF Clare Hall Laboratories, South Mimms, UK); L. Cavalli-Sforza (*D11S16*) (Stanford Univ, California); R. White (*D12S2*) (Howard Hughes Medical Institute, University of Utah); W. Cavenee (*D13S1*, *D13S2*, *D13S4*) (Ludwig Institute for Cancer Research, Montreal). In particular we thank P. Pearson (University of Leiden, The Netherlands) for probe L1.4. We thank Mrs P. Miller for the smooth running of the laboratory.

Received 22 June; accepted 21 July 1987.

- Knudson, A. G. Jr *Proc. natn. Acad. Sci. U.S.A.* **68**, 820-823 (1971).
- Knudson, A. G. *Rev. Genet.* **20**, 231-251 (1986).
- Cavenee, E. K. *et al. Nature* **305**, 779-784 (1983).
- Koufos, A. *et al. Nature* **309**, 170-172 (1984).
- Orkin, S. H., Goldman, D. S. & Sallan, S. E. *Nature* **309**, 172-174 (1984).
- Reeve, A. E. *et al. Nature* **309**, 174-176 (1984).
- Fearon, E. R., Vogelstein, B. & Feinberg, A. P. *Nature* **309**, 176-178 (1984).
- Seizinger, B. R., Martuza, R. L. & Gusella, J. F. *Nature* **322**, 644-647 (1986).
- Hansen, M. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 6216-6220 (1985).
- Koufos, A. *et al. Nature* **316**, 330-334 (1985).
- Bulow, S. *Dan. med. Bull.* **34**, 1-15 (1987).
- Herrera, L., Kakati, S., Gibas, L., Pietrzak, E. & Sandberg, A. A. *Am. J. med. Genet.* **25**, 473-476 (1986).
- Wong, Z., Wong, Z., Wilson, V., Patel, I., Povey, S. & Jeffreys, A. J. *Ann. hum. Genet.* **51**, 269-288 (1987).
- Bodmer, W. F. *et al. Nature* **328**, 614-616 (1987).
- Pearson, P. *Cytogenet. Cell Genet.* **40**, 398 (1985).
- Becher, R., Gibas, Z. & Sandberg, A. A. *Cancer Genet. Cytogenet.* **8**, 329-332 (1983).
- Levin, B. & Reichmann, A. *Cancer Genet. Cytogenet.* **19**, 159-162 (1986).
- Reichmann, A., Martin, P. & Levin, B. *Int. J. Cancer* **28**, 431-440 (1981).
- Paraskeva, C., Paraskeva, C., Buckle, B. G., Sheer, D. & Wigley, C. B. *Int. J. Cancer* **34**, 49-56 (1984).
- Mathew, C. G. P. *et al. Nature* **328**, 524-526 (1987).
- Dukes, C. E. *J. Path. Bact.* **35**, 323-332 (1932).
- Bartels, I., Grzeschik, K.-H., Cooper, D. N. & Schmidtke, J. *Cytogenet. Cell Genet.* **40**, 577 (1985).
- Brennan, J., Spurr, N. K. & Bodmer, W. F. *Nucleic Acids Res.* **14**, 9224 (1986).
- Feder, J. *et al. Am. J. hum. Genet.* **37**, 635-649 (1985).
- Sykes, B. *Disease Markers* **1**, 141-146 (1983).
- Balazs, I., Purrello, M., Kurnit, D. M., Grzeschik, K.-H. & Siniscalco, M. *Somatic Cell molec. Genet.* **10**, 385-397 (1984).
- Cavenee, E. K., Leach, R., Mohandas, T., Pearson, P. & White, R. *Am. J. hum. Genet.* **36**, 10-24 (1984).
- Rider, S. H. *et al. Nucleic Acids Res.* **15**, 863 (1987).
- Baker, D. & White, R. *Cytogenet. Cell Genet.* **32**, 253 (1982).
- Carrel, S., Sordat, B. & Merenda, C. *Cancer Res.* **36**, 3978-3984 (1976).
- Jones, S. L. *Pathology* **14**, 405-408 (1982).
- Drewinko, B., Romsdahl, M. M., Yang, L. Y., Ahearn, M. J., & Trujillo, J. M. *Cancer Res.* **36**, 467-475 (1976).
- Rutzky, L. P. & Siciliani, M. J. *J. natn. Cancer Inst.* (1982).
- Leibovitz, A. *Cancer Res.* **36**, 4562-4567 (1976).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503 (1975).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).

cDNA clones reveal differences between human glial and endothelial cell platelet-derived growth factor A-chains

Benton D. Tong*, David E. Auer*, Michael Jaye†, June M. Kaplow†, George Ricca†, Evelyn McConathy†, William Drohan† & Thomas F. Deuel*‡

* Departments of Medicine and ‡ Biological Chemistry, Jewish Hospital at Washington University Medical Center, St Louis, Missouri 63110, USA

† Department of Molecular Biology, Meloy Laboratories, Inc., Springfield, Virginia 22151, USA

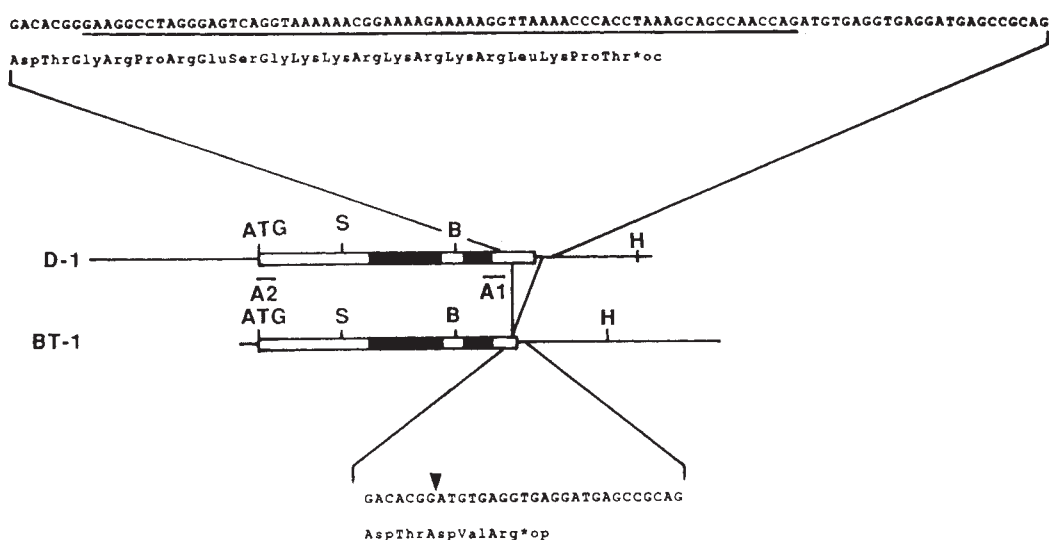
Human platelet-derived growth factor (PDGF) is a potent mitogenic polypeptide¹ which is believed to be a heterodimer of A- and B-chains stabilized by interchain disulphide bonds^{2,3}. The B-chain of PDGF is encoded by the *c-sis* gene, the normal cellular homologue of the transforming gene of the simian sarcoma virus (SSV)⁴⁻⁶. cDNA clones of the B-chain from both normal and transformed cells have mutually consistent DNA sequences^{7,8}. Recently, an A-chain cDNA clone (D-1) was isolated from a transformed human glial cell cDNA library⁹. We report the complete sequence of an A-chain cDNA clone (BT-1) isolated from a normal human umbilical vein endothelial (HUVE) cell cDNA library. BT-1 differs from the sequence of the D-1 clone by a 69 base pair deletion containing the predicted carboxy terminus of the protein. The mRNA levels of the A- and B-chains of PDGF in HUVE cells were analysed and shown to respond differently to the endothelial cell growth factor (ECGF).

Three different cDNA clones were isolated from a λ gt11 HUVE cell-derived cDNA library⁸ using two oligonucleotide probes (A-1 and A-2, Fig. 1), subcloned into the *Eco*RI site of pGEM-2, and sequenced¹⁰. The DNA sequence of the largest cDNA clone (BT-1) begins at nucleotide 351 and continues to about 250 base pairs beyond the 3' end of D-1⁹. The predicted amino acid sequence of BT-1 agreed with previous data obtained from amino acid sequencing, except that Ile 119, Glu 141 and Ser 143 were reported as Val, Arg and Thr respectively⁴. The DNA sequence of all three HUVE cell cDNA clones agreed with the DNA sequence of the transformed glioma cell cDNA clone D-1, except for the deletion of a 69 base pair sequence between D-1 residues 967 and 1,036⁹, located between residues 616 and 617 of BT-1 (Fig. 1). The sequence of BT-1 predicts a protein lacking the 18 C-terminal amino acids found in D-1 and the addition of Asp-Val-Arg (Fig. 1). The domain deleted from D-1 is very basic and includes the sequence Lys, Lys, Arg, Lys, Arg, Lys, Arg which is nearly identical to the nuclear transport signal of the SV40 T-antigen^{11,12}. Both BT-1 and D-1 contain the sequence Val, Arg, Lys, Lys, Pro, Lys which is similar to the nuclear transport signal of the polyoma T-antigen¹³. The overall charge is reduced from +11 (D-1) to +4 (BT-1) and the predicted relative molecular mass of the unprocessed peptide is 22,252.

An oligonucleotide probe within the 69 base pair insert ('insert specific') and a probe of 16 base pairs on each side of the insert ('insert spanning') were used to detect mRNA containing the 69 base pair insert of the D-1 clone in other cell lines. No hybridization of the insert specific probe to A-chain mRNA from two HUVE cell lines or from the human tumour cell lines HOS and A172¹⁴, and U-2 OS¹⁵ was observed under conditions of low stringency. Hybridization of the insert spanning probe identified A-chain mRNA of 2.8, 2.3 and 2.0 kilobases in all five cell lines (data not shown).

Significant differences between the protein products of D-1 and BT-1 were predicted, so BT-1 cDNA was subcloned into an expression vector and transfected into CHO cells. Conditioned media from the BT-1 transfected CHO cells stimulated-

Fig. 1 Comparison of the nucleotide sequence and predicted protein sequence of the C-terminal domains of cDNA clones D-1 and BT-1. A human umbilical vein endothelial cell mRNA cDNA library was constructed in λ gt11 as previously described⁸. Three independent clones hybridized to both the A-1 and A-2 probes. The inserts were isolated and subcloned into the *Eco*RI site of pGem2 (Promega Biotec) and sequenced by the dideoxy method¹⁰ using either the SP6 and T-7 primers (Promega Biotec) or 17–22 base-pair long oligonucleotide primers. The nucleotide sequence and the predicted amino-acid sequence of the C-terminus of both A-chain clones BT-1 (610–640) (complete sequence available upon request) and D-1 (961–1,060)⁹ are shown. The location of the 69 base-pair insert in D-1 is underlined. ∇ Indicates the location of the deletion in BT-1. The open reading frames of both clones are indicated by the boxes and the location of amino acids previously sequenced⁴ are indicated by the filled portions. The relative positions of BT-1 with respect to D-1 is shown. A-1, A-2 indicate locations of the two oligonucleotide probes A-1 and A-2, respectively.



thymidine incorporation into DNA of Swiss 3T3 cells² but conditioned media from control CHO cells lacked growth factor activity. The BT-1 cDNA clone thus encodes an active mitogen.

HUVE cells have been shown to divide and form monolayers in culture when incubated in the presence of ECGF¹⁶. When monolayers are trypsinized and replated without ECGF on to plates retaining residual matrix proteins, they cease to divide and organize into thin, tube-like structures¹⁷. Poly(A)⁺ RNA from endothelial cell monolayers and from endothelial cells in tube-like structures were analysed with probes of the A- and B-chains of PDGF; up to a twofold increase of A-chain mRNA was found in monolayers (Fig. 2a, lane 2) relative to tube forms (Fig. 2a, lane 1). In contrast, B-chain mRNA levels were increased about fivefold (Fig. 2b, lanes 1, 2)¹⁸. The increase in levels of B-chain mRNA were repeatedly in marked excess of increases in the A-chain mRNA levels expressed in the monolayer cells. Figure 2 also presents Northern blot analysis of the A172 cell line poly(A)⁺ RNA. The A172 cell contains significant levels of both A- and B-chain transcripts (Fig. 2a, b, lanes 3). The relatively high levels of A-chain mRNA in A172 cells in comparison to HUVE cells were less striking when total RNA from both cell lines was examined. Southern blot analysis with a nick-translated A-chain cDNA probe demonstrated quantitatively identical results with DNA from A172 cells and from HUVE cells, therefore no evidence for gene amplification in A172 cells was observed (data not shown).

Previously, cDNA clones for the B-chain of PDGF have been isolated and shown to have no significant differences from clones isolated from transformed cells^{7,8,19,20}. In contrast, the data presented in this report suggest that an A-chain mRNA represented by the D-1 clone (containing the 69 base pair insert) previously isolated and sequenced from transformed glial cells⁹ was not present in any of three independent cDNA clones from a HUVE cell cDNA library nor was it found in RNA from HUVE cells using an insert specific probe in Northern blot analysis. A second clone in the transformed glial cell cDNA library lacked the same 69 base pair sequence as the three cDNA clones from HUVE cells⁹. The D-1 clone is likely to be the product of alternative splicing in the transformed glial cell line, although a gene polymorphism cannot be ruled out. In the three other tumour cell lines which express significant levels of A-chain mRNA, we did not detect the D-1 69 base pair insert, suggesting the 69 base pair containing clone cannot alone account for transformation of the glial cell line.

The protein predicted from the BT-1 cDNA clone is significantly different from the D-1 clone in being substantially less basic (from +11 to +4). The protein product of the BT-1 cDNA clone from transfected CHO cells is, however, mitogenically active. The presence of a signal peptide suggests that predicted proteins of both D-1 and BT-1 would be processed through the endoplasmic reticulum/Golgi apparatus and that the nuclear transport signals would not normally be of importance. The data also suggest that in endothelial cells, the levels of A-chain and B-chain mRNA are regulated differently. This leaves unanswered how the A- and B-chains may be coordinately expressed in megakaryocytes to form the apparent heterodimeric

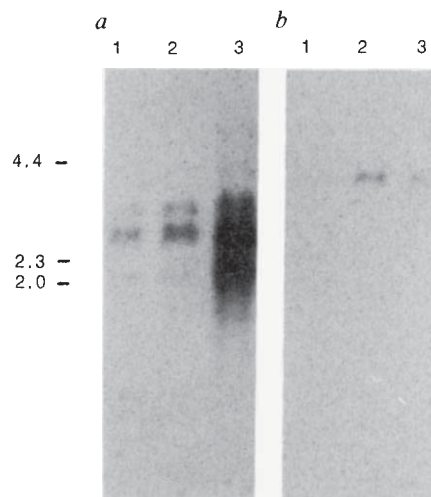


Fig. 2 Modulation of PDGF A- and B-chain mRNA in endothelial cells. RNA was isolated as previously described¹⁸ and electrophoresed in a 1% agarose gel containing formaldehyde and transferred to nitrocellulose⁸. The filter was first hybridized as previously described⁸ to nick-translated BT-1 (a) and the probe removed from the filter using 0.1 M NaOH 0.5 M NaCl, and then rehybridized to nick-translated *v-sis* (b). Filters were exposed to XAR5 film overnight at -70°C using a Dupont Cronex intensifying screen. *Hind*III digested λ DNA was used as size markers. Five μg of poly(A)⁺ RNA from tube form (lane 1), monolayer (lane 2) human umbilical vein endothelial cells or A172 cells (lane 3) was used. No hybridization was observed to either tRNA or poly(A)⁺ RNA from rat promegakaryoblast cells²¹.

structure of human PDGF, and opens the possibility of potentially independent roles of A- and B-chain homodimers in endothelial cell function. The expression of the A-chain in endothelial cells is potentially significant and implies an important role for the endothelial cell in the normal and perhaps abnormal paracrine stimulation of cells of the blood vessel wall.

We thank R. Kawahara and F. D. Porter for stimulating discussions, C.-H. Heldin for A-chain clone D-1, L. Chasin for CHO cells, D. Hurwitz for expression vector construction, and B. B. Kennedy and R. Mudd for technical assistance. B.D.T. was supported in part by a Public Health Service training grant from the National Institutes of Health. The work was supported by Public Health Service grants and by a grant from the Monsanto Corporation to T.F.D.

Received 9 March; accepted 12 June 1987.

1. Ross, R. & Vogel, A. *Cell* **14**, 203-210 (1978).
2. Deuel, T. F. *et al. J. biol. Chem.* **256**, 8896-8899 (1981).
3. Johnsson, A., Heldin, C.-H., Westermark, B. & Wasteson, A. *Biochem. biophys. Res. Commun.* **104**, 66-74 (1982).
4. Waterfield, M. D. *et al. Nature* **304**, 35-39 (1983).
5. Doolittle, R. F. *et al. Science* **221**, 275-277 (1983).
6. Johnsson, A. *et al. EMBO J.* **3**, 921-928 (1984).
7. Ratner, L., Josephs, S. F., Jarrett, R., Reitz, M. S. Jr & Wong-Staal, F. *Nucleic Acid Res.* **13**, 5007-5018 (1985).
8. Tong, B. D. *et al. Molec. cell. Biol.* **6**, 3018-3022 (1986).
9. Betsholtz, C. *et al. Nature* **320**, 695-699 (1986).
10. Chen, E. Y. & Seeburg, P. H. *DNA* **4**, 165-170 (1985).
11. Lanford, R. E. & Butel, J. S. *Cell* **37**, 801-813 (1984).
12. Kalderon, D., Richardson, W. D., Markham, A. F. & Smith, A. E. *Nature* **311**, 33-38 (1984).
13. Richardson, W. D., Roberts, B. L. & Smith, A. E. *Cell* **44**, 77-85 (1986).
14. Eva, A. *et al. Nature* **295**, 116-119 (1982).
15. Graves, D. T. *et al. Science* **226**, 972-974 (1984).
16. Maciag, T., Mehlman, T., Friesel, R. & Schreiber, A. B. *Science* **225**, 932-935 (1984).
17. Maciag, T., Kadish, J., Wilkens, L., Stemerman, M. B. & Weinstein, R. J. *Cell Biol.* **94**, 511-520 (1982).
18. Jaye, M. *et al. Science* **228**, 883-885 (1985).
19. Collins, T., Ginsberg, D., Boss, J. M., Orkin, S. H. & Pober, J. S. *Nature* **316**, 748-750 (1985).
20. Rao, C. D., Igarashi, H., Chiu, I.-M., Robbins, K. C. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2392-2396 (1986).
21. Weinstein, R., Stemerman, M. B., MacIntyre, D. E., Steinberg, H. N. & Maciag, T. *Blood* **58**, 110-121 (1981).

Alternative RNA splicing affects function of encoded platelet-derived growth factor A chain

Tucker Collins*, David T. Bonthron†‡
& Stuart H. Orkin†‡

* Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115, USA

† Division of Hematology/Oncology, Children's Hospital and the Dana Farber Cancer Institute; Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Howard Hughes Medical Institute, 300 Longwood Avenue, Boston, Massachusetts 02115, USA

Platelet-derived growth factor (PDGF) is a basic protein of relative molecular mass 30,000 (M_r 30K) composed of two polypeptide chains, designated PDGF A and PDGF B^{1,2}. The B-chain is encoded by the *c-sis* gene, the cellular counterpart of the simian sarcoma virus transforming gene *v-sis*^{1,3,4}. The PDGF A-chain cDNA clones recently isolated and sequenced from a transformed human clonal glioma cell line represent at least two alternatively spliced transcript species differing by 69 base pairs at the C-terminus⁵. Here we demonstrate that the normal human umbilical vein endothelial cell (EC) A chain precursor lacks the 15 carboxy-terminal, highly basic amino acids encoded by the larger tumour cell cDNA. Surprisingly, culture media from monkey kidney cells (COS) transfected with the endothelial cDNA clone contained much less mitogenic activity than media from cells transfected

with the longer tumour cell-derived A-chain cDNA. This functional difference appeared to be due to inefficient assembly or secretion of the recombinant endothelial-type growth factor. This suggests that some transformed cells may use alternative RNA splicing to modify normal growth factors and by so doing increase the efficiency of mitogen assembly or secretion.

Human umbilical vein endothelial cells (EC) express both PDGF A- and B-chain mRNA^{6,7}. Several A chain cDNA clones were isolated from EC cDNA libraries and the longest (1.75 kilobases (kb)) was completely sequenced. The PDGF A-chain precursor open reading frame (196 amino acids, M_r 22,251) was flanked by 5' and 3' non-coding regions of 585 and 589 base pairs (bp) respectively. The coding region is identical to a previously reported glioma cDNA clone⁵ from amino acid residue 1 to 193, at which point the sequences diverge. At the nucleotide level, the sequences are identical to bp 1,165; after an insertion of 69 bp in the glioma sequence (Fig. 1), the two clones resume identity. The difference in transcript structure shortens the predicted endothelial precursor by 15 amino acids, eliminating a highly basic carboxy-terminal region from the glioma-derived precursor. Sequence analysis of three additional clones, obtained from an independent EC cDNA library prepared by random hexanucleotide priming, revealed a carboxy-terminal structure identical to that of the clone obtained from the oligo(dT) primed library.

The possibility that the longer A-chain transcript (glioma-like species) might be present in the EC as a minor species was tested by solution hybridization and nuclease protection assays. An antisense RNA probe corresponding to the EC carboxy-terminus (Fig. 2a) yielded a single protected fragment of 474 bp; no smaller species were detectable (Fig. 2b). With a probe corresponding to the glioma A-chain carboxy-terminus, fragments of 250 bp and 223 bp were protected. No species corresponding to the glioma cDNA (a protected fragment of 542 bp) was identified in RNA from either EC or two osteosarcoma cell lines (Fig. 2c). Thus the shorter A-chain PDGF mRNA is the predominant, if not sole, transcript in these cells. The EC and the clonal glioma A chain species arise by alternative splicing since a single exon of the A-chain gene encodes the 69 bp inserted segment (D.T.B., S.H.O. and T.C., manuscript submitted).

Assembly of functional PDGF involves dimerization as well as amino- and carboxy-terminal proteolysis (reviewed in ref. 8). To determine whether the structural differences in the EC and glioma PDGF A-chain cDNAs have functional consequences, each A-chain open reading frame was inserted into an SV-40 based expression vector, and transfected into COS-1⁹ cells. These do not contain either A- or B-chain mRNA, or produce a mitogen which stimulates quiescent 3T3 cells (Figs 3, 4). Conditioned media from transiently-transfected COS cells were transferred to quiescent mouse 3T3 cells and the induction of proliferation measured by a ³H-thymidine incorporation assay¹⁰. Medium from cells transfected with the glioma PDGF A-chain markedly stimulated ³H-thymidine uptake, whereas medium from control cells failed to induce proliferation (Fig. 3). Conditioned media from several transfections with the EC-derived A-chain expression construct gave low but detectable stimulation of ³H-thymidine uptake (Fig. 3). A similar difference in activity between glioma and EC expression constructs was seen when NIH 3T3 cells were transfected with the PDGF A-chain constructs and the mitogenic activity in the conditioned medium was measured (data not shown). This suggests that the low mitogenic activity directed by the EC cDNA in COS cells is not specific to the COS-1 cell, but is an inherent property of the EC A-chain precursor. Northern blots of poly(A⁺) RNA indicated that similar transcript levels were present after transfection with both types of A-chain construct (Fig. 4a). Also, COS cells transfected in parallel were metabolically labelled with ³⁵S-cysteine. Figure 4b,c shows SDS polyacrylamide gel electrophoresis (SDS-PAGE) of total labelled protein from the

Fig. 1 Comparison of the PDGF precursor sequences. Top, schematic representation of the EC and the glioma mRNAs. The shaded area bounded by triangles denotes the 69-bp sequence in the glioma cDNA clone absent from the EC cDNA clone. Arrows, boundaries of the N- and C-terminal propeptides. Middle, sequences of the glioma (clone D1)⁵ and the endothelial (clone 1-1) cDNAs from positions 964 to 1,042 and 1,160 to 1,176, respectively. In-frame stop codons completing the A-chain open reading frame are underlined. A *Stu*I site unique to the glioma cDNA is marked by a dashed line. The last two lines align the C-terminal amino-acid sequences of the precursors.

Methods. HUVE culture and constructure of cDNA libraries are described in detail elsewhere^{11,12}. Recombinant λ gt11 phage clones were isolated by screening with an 844-bp *Sst*II-*Hind*III fragment of the glioma cDNA clone⁵, radiolabelled by nick-translation¹³, or by oligonucleotide primer labelling¹⁴. cDNA inserts were subcloned into pUC-19¹⁵. Both strands of the 1-1 cDNA clone from the oligo(dT) library were completely sequenced using the Sanger dideoxy method and the Maxam-Gilbert chemical degradation methods^{16,17}.

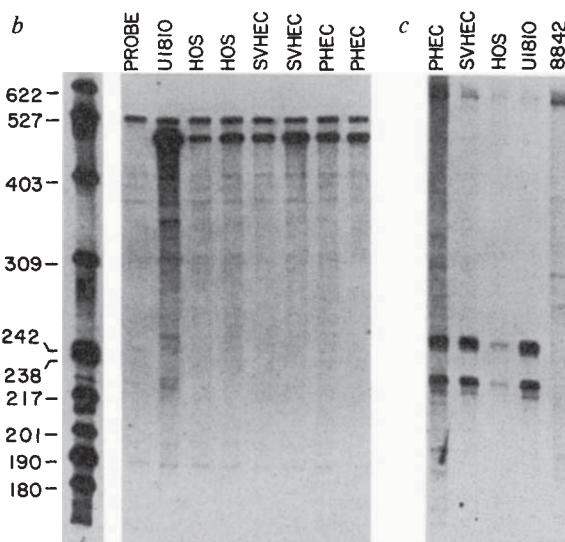
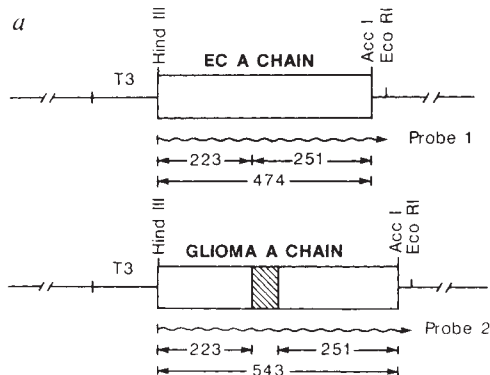
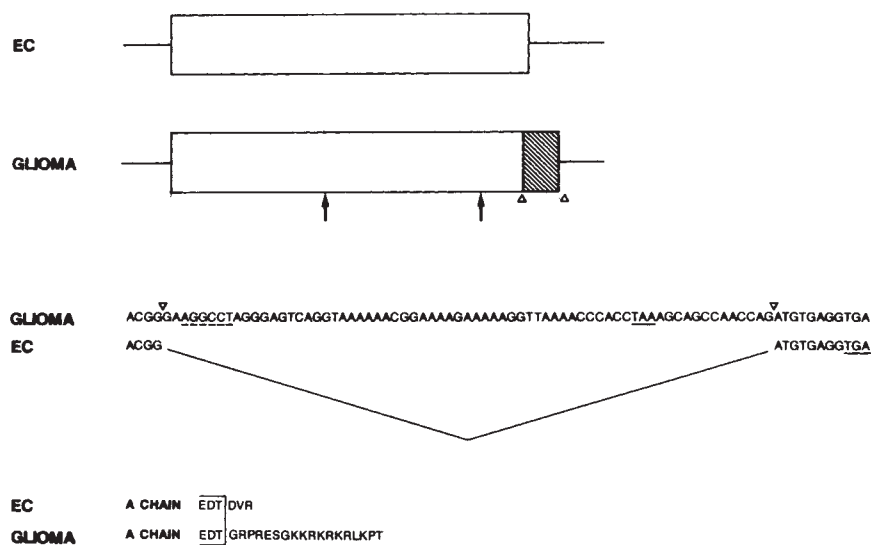


Fig. 2 Identification of the PDGF A-chain transcript present in EC by RNase mapping. RNAs were analysed by RNase protection assays using probes containing the endothelial (probe 1) or glioma (probe 2) A chain carboxy-terminus and adjacent 3' untranslated region. **a**, diagram of two antisense probes and the expected protected fragments; **b**, RNase protection assay with probe 1; **c**, RNase protection assay with probe 2; M, markers (*Hpa*II-digested pBR322, 3' end labelled with ³²P); PROBE, antisense RNA probe; PHEC, passaged human EC; SVHEC, SV40 transformed EC; HOS, a human osteosarcoma cell line; U1810, a second human osteosarcoma cell line; 8842, a third human osteosarcoma cell line which does not express PDGF A-chain.

Methods. *Acc*I-*Hind*III fragments of the EC and glioma cDNA clones were subcloned into the multiple cloning site of Bluescribe M13+ (Stratagene). The constructs were made linear with *Eco*RI and antisense transcripts were synthesized *in vitro* using T3 RNA polymerase and ³²P-rCTP (800 Ci/mMole, Amersham) according to the manufacturer's instructions. Cytoplasmic RNA was isolated as described previously^{6,7}. Samples of 10 and 25 μ g (**b**) or 25 μ g (**c**) of cytoplasmic RNA (or 5 μ g of poly(A⁺) RNA-U1810), were hybridized at 45 °C overnight with the labelled antisense transcripts and the hybrids were digested for 60 min at 30 °C with RNase A and T₁ as described¹⁸. The protected fragments were separated on 8% polyacrylamide/urea gels.

conditioned media of the transfectants; Fig. 4d shows material reactive with an antihuman PDGF heteroserum. A major new species of about 32K was seen in the medium from cells transfected with the glioma A-chain cDNA (Fig. 4b); its identity was confirmed by immunoprecipitation with an anti-PDGF antibody (Fig. 4d). In samples reduced with mercaptoethanol, a major 18.5K and a minor 26K species were identified (Fig. 4c). These are the expected sizes for cleaved A-chain and uncleaved pro-A

PDGF, respectively. Cleavage of pro-A chain in COS cells is not complete before secretion; however, dimerization appears essentially complete, as no monomer-sized species were seen before reduction. This implies that in COS cells A chain dimerization precedes pro-peptide cleavage. The absence of similar proteins in the conditioned medium of the EC A-chain cDNA transfectants confirmed that the very low mitogenic activity directed by this clone could largely be ascribed to a failure of

Fig. 3 Transient expression of recombinant PDGF A-chain. Mitogenic activity of conditioned medium from COS-1 cells transfected with a plasmid containing the glioma A-chain cDNA-pSV2AC-7 (●—●), a control plasmid-pSV2 (●...●), or two clones of a plasmid containing the EC A-chain cDNA-pSV2AC-15, 17 (○—○). Error bars are omitted where smaller than the symbol size.

Methods. A *Hind*III linker was attached at the *Sac*II site of both the EC and the glioma cDNA clones. *Hind*III fragments (*Sac*II-*Hind*III of the original sequences) were inserted into a modified plasmid pSV2¹⁹. Each 100 mm dish of subconfluent COS-1 cells was transfected with 20 µg plasmid using a modified calcium phosphate procedure²⁰. Two days after transfection, the medium was changed to 5 ml of serum-free Dulbecco's modified Eagle medium (DME). The conditioned serum-free medium was collected 48 h later and mitogenic activity was measured by a modified version of an assay described previously¹⁰. NIH 3T3 cells were seeded at 10⁴ cells per 0.28-cm² well, grown to confluence (3 days) in DME containing 10% fetal calf serum, rinsed and incubated overnight in DME containing 2.5% platelet-poor plasma²¹. The quiescent cells were incubated for 20 h with 0.2-ml test samples, containing 2.5% platelet-poor plasma, and then labelled for 90 min in 0.1 ml DME/2.5% platelet-poor plasma containing 1 µCi ³H-thymidine (5Ci mmole⁻¹, Amersham). Results are shown as c.p.m. (±s.e.m.) incorporated above background into 5% trichloroacetic acid-insoluble material.

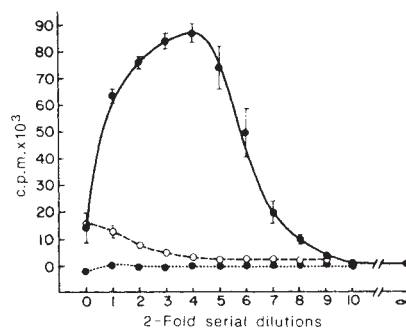
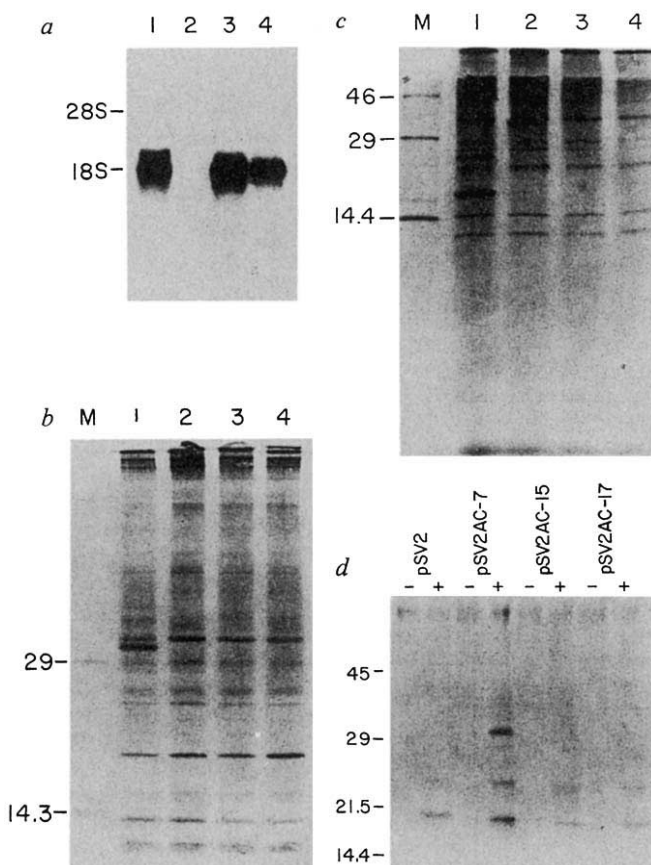


Fig. 4 Analysis of COS cells transfected with PDGF A-chain expression vectors. *a*, Northern analysis of PDGF A-chain mRNA made by COS-1 cells transfected with the A-chain expression constructs. Lane 1, pSV2AC-7; 2, pSV2; 3, pSV2AC-15; 4, pSV2AC-17. *b*, *c*, Metabolically labelled proteins in conditioned media of cells transfected with pSV2AC-7 (lane 1), pSV2 (lane 2), pSV2AC-15 (lane 3) or pSV2AC-17 (lane 4). SDS-PAGE²² was performed before (*b*) or after (*c*) reduction with β -mercaptoethanol. *d*, Immune-precipitated labelled proteins from COS-1 cells, transfected as above and analysed in the absence of reducing agents. Lanes marked (–) are precipitations with normal goat serum and lanes marked (+) are precipitations with goat anti-human PDGF; the A-chain expression constructs used are indicated at the top of the gel.

Methods. Cytoplasmic RNA was isolated by detergent lysis and proteinase K digestion and DNase-I treated. Poly(A⁺) RNA was selected by batch absorption on oligo(dT) cellulose¹³. Poly(A⁺) RNA (~1 µg per lane) was probed with the *Hind*III cDNA fragment from pSV2AC-15. Cells were metabolically labelled for 5 h in cysteine-free MEM containing 0.4 mCi ml⁻¹ ³⁵S-cysteine (600 Ci mmol⁻¹, Amersham). After a chase period of 14 h, the labelled media were dialysed against 10 mM Tris-Cl, pH 6.8, lyophilized, dissolved in water, and transferred to standard Laemmli sample buffer²². For immunoprecipitation, the same labelled media were made 25 mM Tris Cl, pH 7.5, 25 mM NaCl, 0.1 mM PMSF, 0.05% NP-40 before preclearing with normal goat serum and rabbit anti-goat Sepharose beads (Miles Scientific). Samples were then incubated with either goat anti-human PDGF (Collaborative Research) or normal goat serum for 12 h at 4°C. Washed rabbit anti-goat Sepharose beads were added and incubated for 2 h at 4°C before washing twice with the above buffer containing 0.5 M NaCl. The immune complexes were eluted with standard Laemmli sample buffer²².



the COS cells to process, assemble or secrete an active mitogen from the shorter (EC) form of the PDGF A-chain.

The two alternatively spliced mRNAs encode A-chain precursors which differ in their abilities to assemble into mature secreted mitogens. A possible implication of this finding is that the normal EC type of A-chain precursor may require dimerization with a B-chain precursor to allow efficient processing and secretion. The increased mitogen production seen in some pathological settings⁸, however, might coincide with activation of an abnormal splicing pattern (longer mRNA type), allowing secretion of an active A-chain homodimer. Changes in the A-chain splicing pattern may accompany normal embryogenesis and development. Second, the heterogeneity of the growth factor¹ could result not only from carboxy-terminal proteolysis but

also from alternative A-chain mRNA splicing. Such heterogeneity could affect precursor stability or receptor affinity. Third, some transformed cells may use alternative splicing to modify a growth factor and generate a mitogen which, because of more efficient assembly or processing, may be an important component of the transformed phenotype.

We thank C. Betsholtz for providing the glioma PDGF A-chain cDNA probe (clone D1), C.-H. Heldin for helpful comments and support; M. Gimbrone for endothelial cultures, J. Pober for helpful discussion and R. Cotran and J. L. Strominger for their support. This work was partly supported by grants from the NIH to T.C. and S.H.O.; T.C. is a PEW Scholar in the Biomedical Sciences; S.H.O. is an Investigator of the Howard Hughes Medical Institute.

Received 25 March; accepted 1 June 1987.

1. Johnsson, A. *et al.* *EMBO J.* **3**, 921-928 (1984).
2. Deuel, T. F. *et al.* *J. biol. Chem.* **256**, 8896-8899 (1981).
3. Waterfield, M. D. *et al.* *Nature* **304**, 35-39 (1983).
4. Doolittle, R. F. *et al.* *Science* **221**, 275-277 (1983).
5. Betscholtz, C. *et al.* *Nature* **320**, 695-699 (1986).
6. Collins, T., Ginsburg, D., Boss, J. M., Orkin, S. H. & Pober, J. S. *Nature* **316**, 748-750 (1985).
7. Collins, T. *et al.* *Am. J. Pathol.* **126**, 7-12 (1987).
8. Ross, R., Raines, E. W. & Bowen-Pope, D. F. *Cell* **46**, 155-169 (1986).
9. Gluzman, Y. *Cell* **23**, 175-182 (1981).
10. Raines, E. W. & Ross, R. *Meth. Enzym.* **109**, 749-773 (1985).
11. Gimbrone, M. A. Jr *Prog. Hemostasis. Thromb.* **3**, 1-28 (1976).
12. Ginsburg, D. *et al.* *Science* **228**, 1401-1406 (1985).
13. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
14. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
15. Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103-119 (1985).
16. Sanger, F., Nicklen, S., & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
17. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-553 (1980).
18. Melton, D. A. *et al.* *Nucleic Acids Res.* **12**, 7035-7056 (1984).
19. Orkin, S. H., Goff, S. C., Kelley, W. N. & Daddona, P. E. *Molec. cell. Biol.* **5**, 762-767 (1985).
20. Chen, C. & Okayama, H. *Molec. cell. Biol.* (in the press).
21. Pledger, W. J., Stiles, C. D., Antoniadis, H. N. & Scherr, C. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4481-4485 (1977).
22. Laemmli, U. K. *Nature* **227**, 680-685 (1970).

Antiglucocorticosteroid effects suggest why steroid hormone is required for receptors to bind DNA *in vivo* but not *in vitro*

André Groyer, Ghislaine Schweizer-Groyer,
Françoise Cadepond, Marcel Mariller
& Etienne-Emile Baulieu

INSERM U 33 and Faculté de Médecine, 94275 Bicêtre Cedex, France

Sequence-specific interaction between steroid hormone receptors (R) and DNA hormone-responsive elements (HRE) takes place *in vitro* irrespective of the presence of hormone and even when R is liganded with an antagonist^{1,2}. *In vivo*, in contrast, the presence of hormone is mandatory for glucocorticosteroid (G) receptor-HRE interaction to occur and no HRE occupancy is detected in the presence of an antagonist³. One possible explanation is that *in vivo* R is originally complexed with a protein that prevents its binding to target HREs. The hormone would then induce the dissociation of the oligomer, thus unmasking the functional DNA binding domain of the receptor. The unliganded, non DNA-binding 8S-form of the chick GR is a hetero-oligomer including the relative molecular mass (M_r) 94,000 steroid-binding unit (4S-GR), and the non-steroid-binding, non-DNA-binding 90,000 protein⁴ common to all classes of 8S-R⁵ and identified as heat-shock protein (hsp 90)⁶⁻⁸. We report here that triamcinolone acetonide (TA) promotes the transformation of 8S-GR to 4S-GR complexes both in explants and in cell-free conditions and that the high-affinity antiglycorticosteroid RU 486⁹ stabilizes the 8S-GR, as assessed by gradient sedimentation and HPLC. However, *in vitro* TA- and RU 486- 4S-GR showed comparable DNA-binding activity. These results suggest that the lack of affinity for DNA of the 8S form of GR may be attributable *in vivo* to the interaction of the 4S-GR protein with hsp 90, and that hormone binding might trigger a conformational change which results in the release of active 4S-GR.

In the chick oviduct, the glucocorticosteroid-induced increase in egg-white protein gene transcription proceeds through binding to the glucocorticosteroid receptor (GR) and is inhibited by RU 486¹⁰⁻¹². To understand the mechanism of RU 486 action, we have examined its effects on GR nuclear binding in whole cells. When explants of magnum were incubated with RU 486, which does not bind to the chick progesterone receptor⁴, almost no RU 486-GR complexes were found in the purified nuclei.

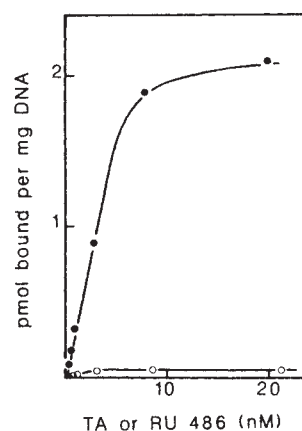


Fig. 1 Whole cell nuclear binding of triamcinolone acetonide and RU 486. Saturable binding of GR to nuclei was measured after incubation of magnum tissue (~1 mm³ pieces) with various concentrations of TA (0.23-20 nM) (●), or RU 486 (0.24-21 nM) (○). TA was preferred to dexamethasone as it binds exclusively to GR, and not to the cytosoluble transcortin-like protein, and because TA-GR is more stable than dexamethasone-GR⁴.

Methods. Oviducts were excised from diethylstilboestrol-primed chicks, 4 days after oestrogen withdrawal. Chopped magnum tissue (100 mg) was incubated at 37 °C for 1 h in RPMI 1640 medium (1 ml), buffered with 25 mM Hepes with ³H-TA or ³H-RU 486, non-saturable binding being assessed in parallel incubations carried out in the presence of unlabelled steroid (2 μM). Tissue fragments were centrifuged, washed three times with cold PBS (4 ml) and manually homogenized in 10 volumes of TKSS buffer (15 mM Tris-HCl, pH 7.4-7.5, 60 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 15 mM NaCl, 15 mM 2-mercaptoethanol) containing 0.34 M sucrose, 2 mM EDTA, 0.5 mM EGTA and 0.2 mM PMSF. Nuclei were prepared as previously described^{11,12} using a modified procedure²³. The final nuclear pellet was resuspended in 0.4 ml of TKSS buffer containing 0.34 M sucrose. Replicate 100-μl aliquots were counted or used for DNA estimation.

This contrasts with the nuclear localization of ~80% of intracellular TA-GR (2.58 ± 0.36 pmol per mg DNA) after incubation with TA (Fig. 1). It was found that nuclear binding of TA was inhibited by RU 486 in a dose-dependent manner (data not shown). These results, together with the data reported so far¹³⁻¹⁶, indicate that the amount of nuclear-bound RU 486-GR ranges from 5-10% (purified nuclei^{13,14}) to ~40% (crude nuclear pellets^{15,16}) of that of agonist-GR. The latter is likely to be an overestimate, arising from loosely-bound nuclear complexes¹⁷ becoming trapped in cell debris after homogenization. The low affinity of non-transformed 8S-GR and the high affinity of the transformed 4S-GR for nuclei and DNA^{18,19} led us to investigate the influence of RU 486 binding on the transformation process and to relate the size of RU 486-GR with its affinity for DNA.

Analysis on glycerol gradients showed that untransformed TA-GR sedimented at 8.5 ± 0.5 S ($n = 9$) in low salt buffer (Fig. 2a), irrespective of the presence of sodium molybdate (50 nM)⁴. After transformation (1 h at 0 °C with 0.3 M KCl in absence of molybdate), most TA-GR was shifted to 4.4 ± 0.3 S ($n = 5$). After this treatment, cytosol was incubated with BF₄ monoclonal antibody (BF₄ recognizes hsp 90 but not the transformed 4S steroid-binding unit⁴⁻⁶), and the small residual 8.5 S peak was displaced to 10.6 S (Fig. 2a). In contrast, high-salt treatment did not affect the sedimentation coefficient (8.7 ± 0.6 S; $n = 12$) of RU 486-GR (Fig. 2b). Furthermore, irrespective of whether the cytosol had been exposed to salt, BF₄ interacted with RU 486-GR to sediment at 11.3 ± 0.5 S ($n = 4$). Similar results were obtained when 8S-GR was transformed at 25 °C for 1 h and the size estimated by HPLC: the untransformed TA-GR and RU 486-GR had a Stokes radius (R_s) of 6.7 ± 0.1 nm ($n = 11$ and 6 respectively) and the R_s of transformed TA-GR was 4.9 ± 0.1 nm

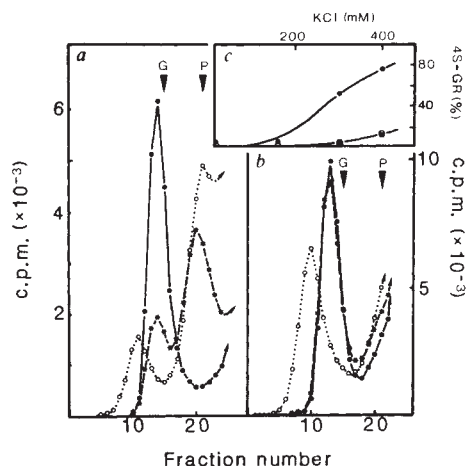


Fig. 2 Sedimentation on glycerol gradients of: *a*, ^3H -triamcinolone acetonide-GR; and *b*, ^3H -RU 486-GR before (●) and after (○,*) high-salt transformation. In the latter case, the sedimentation patterns of TA- and RU 486-GR were analysed with (○) and without (*) previous incubation with BF_4 monoclonal antibody. Glucose oxidase (G) and peroxidase (P) are internal markers. *c*, % of 4S-GR obtained after incubation of TA-GR (●), RU 486-GR (○) and unliganded GR (■) with increasing KCl concentrations (0–0.4 M).

Methods. Cytosol was prepared in GTED buffer (15% glycerol, 20 mM Tris-HCl, pH 7.4, 1.5 mM EDTA, 1 mM dithioerythritol) (4 ml g^{-1}), and duplicate 500 μl samples were incubated with ^3H -TA or ^3H -RU 486 (20 nM). Incubation was for 4 h (●) or 3 h followed by 1 h with 0.3 M KCl present (*). After 4 h samples were made 50 mM in sodium molybdate and unbound steroid was removed by treatment for 10 min with dextran-coated charcoal (0.5% Norit A, 0.05% Dextran T70). For testing of immunorecognition of TA-GR and RU 486-GR, a further 5 h incubation was performed with BF_4 monoclonal antibody (35 μl of labelled cytosol with 75 μl of purified BF_4 in a total volume of 200 μl). Incubations in *c* were for 1 h with increasing concentrations of KCl (0–0.4 M), before or after labelling with ^3H -TA or ^3H -RU 486 for 3 h. Samples (200 μl) were spun at 50,000 r.p.m. through 10–35% glycerol gradients in TED buffer made 50 mM in molybdate, for 17 h (Beckman SW60 rotor). The saturable binding measured in the 4S region of the gradient was expressed as % of total saturable (8S+4S) binding. Inactivation of GR was always lower than 25% in these experiments. Each point represents the mean of two determinations performed with different cytosol preparations. All operations were carried out at 0°C.

($n = 6$), that of RU 486-GR remaining unchanged at $6.7 \pm 0.1 \text{ nm}$, $n = 8$ (data not shown).

Transformation was checked after exposure of 8S-GR to different concentrations of KCl (0–0.4 M) at 0°C. A 50% conversion of 8S to 4S was obtained at 0.3 M KCl for TA-GR, while <10% transition of 8S to 4S occurred for RU 486-GR or in the absence of steroid ligand (Fig. 2*c*). Temperature variation (15–35°C) gave a 50% conversion of TA-GR at 25°C when 95% of RU 486-GR remained in the 8S form (data not shown).

Whole cell studies were performed to check whether RU 486 impairs GR transformation *in vivo*. When explants of magnum were incubated at 37°C for 1 h with ^3H -TA, then homogenized in low salt buffer, ~20% of total GR was recovered in the cytosol, almost all (88%) being 4S (Fig. 3*a*). When RU 486 was used in the incubation, ~94% of total GR was cytosoluble, 59.3% being in 8S form (Fig. 3*b*). We conclude that in whole cells the transformation of RU 486-8S-GR was impaired compared to TA-8S-GR (~40 and ~98% of total GR, respectively). The presence of some 4S-GR in the cytosol during these experiments, lower in the case of TA than for RU 486, cannot easily be explained but does not affect the overall trend of the results.

The DNA binding activity of the TA- and RU 486-GR was tested using DNA-cellulose. Untransformed or preheated

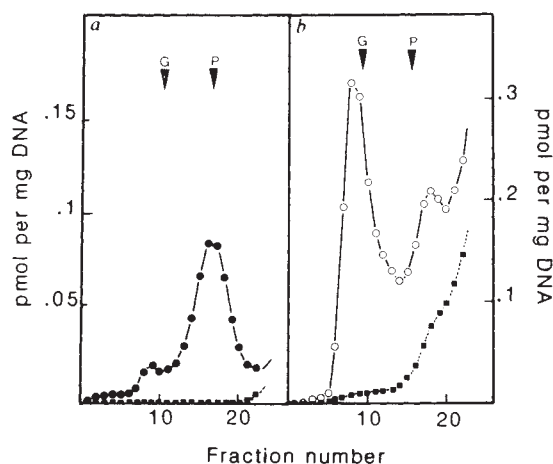


Fig. 3 Sedimentation on glycerol gradients of: *a*, the cytosoluble ^3H -triamcinolone acetonide-GR (●) and *b*, ^3H -RU 486-GR (○) after tissue mince incubation. (■) Non-saturable binding was estimated in parallel incubations with 2 μM competing unlabelled steroids to assess non-saturable binding. G and P markers as in Fig. 2.

Methods. Explants were prepared and incubated as described in Fig. 1, except that a single concentration (20 nM) was used for both ^3H -TA and ^3H -RU 486. Incubated tissue was washed with PBS (4 ml), homogenized in GTED-M50 buffer (GTED buffer with 50 mM sodium molybdate, 1 ml g^{-1}) containing unlabelled TA or RU 486 (2 μM), and diluted twice with TED-M50 buffer. Homogenates were spun at 3,600g for 10 min and the supernatant treated with dextran-coated charcoal for 15 min and further centrifuged at 40,000 r.p.m. for 1 h (Beckman 50 Ti rotor). 200 μl of the high speed supernatants were run on 10–35% glycerol gradients (see legend to Fig. 2).

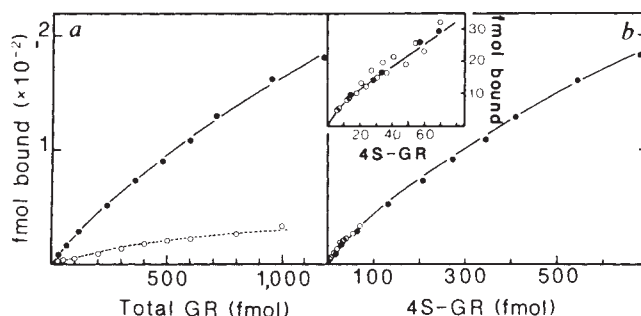


Fig. 4 DNA-cellulose binding of triamcinolone acetonide (●) and RU 486-GR (○). *a*, Variation of binding of agonist- and antagonist-GR to DNA-cellulose with GR input. *b*, DNA-cellulose binding as a function of amount of 4S-GR. Inset shows duplicate results on a different scale.

Methods. Cytosol in GTED buffer (3 ml g^{-1}) (1 ml) was incubated with ^3H -TA or ^3H -RU 486 (20 nM) for 3 h on ice, then for 1 h at 25°C. All samples were adjusted to 50 mM in sodium molybdate and unbound steroid removed by dextran-coated charcoal (see Fig. 2 legend). Duplicate aliquots of labelled cytosol (5–200 μl) were incubated at 0°C for 45 min with 200 μl DNA-cellulose²⁴ slurry (1.35 mg of calf thymus DNA per ml of packed gel) with gentle shaking. The gel was then washed 3 times with GTED-M50 and counted. Data are corrected for non saturable binding.

(25°C, 1 h) cytosoluble TA- and RU 486-GR did not bind significantly to the resin in the presence of sodium molybdate. Preheating without molybdate promotes the binding of TA- and RU 486-GR to DNA-cellulose, being five times larger for the former than that of RU 486-GR relative to total receptor input (Fig. 4*a*). However, DNA-cellulose binding of TA- and RU 486-GR was indistinguishable when replotted as function of the amount of 4S receptor added (representing ~57% of TA-GR

and only ~7% of RU 486-GR) (Fig. 4b). Two series of complementary experiments confirmed this observation. First, when transformed (4S) TA- and RU 486-GR were isolated on glycerol gradients, their binding to DNA-cellulose was identical (data not shown). Second, when DNA-cellulose bound TA- and RU 486-GR were eluted with increasing concentrations of KCl in the washing buffer (20–200 mM)²⁰, the same patterns of elution were obtained (binding was decreased by 50% at 133.2 ± 10.4 and 128.2 ± 9.5 mM for TA- and RU 486-GR, respectively). Our results agree with recent findings indicating that agonist-, antagonist- and ligand-free GR and progesterone R, bind to HREs in mouse mammary tumour virus long terminal repeat (MMTV-LTR) DNA and the uteroglobin gene promoter, respectively (refs 1 and 2 and A.G., G.S.G., F.C., unpublished experiments). Moreover, our study explains DNA-binding data showing that RU 486-GR is less transformed than TA-GR and behaves in some aspects like an nt^- agonist-GR complex^{16,21}.

Even if alternative explanations may be presented²², our data may reconcile different results obtained for DNA interaction of RU 486-GR *in vivo*³ and *in vitro*^{1,2}. Indeed, the limited ability of the 8S, hetero-oligomeric, RU 486-GR to undergo subunit dissociation supports the proposal that, *in vivo*, the antagonistic activity of RU 486 is mainly attributable to the stabilization of the non-transformed form of GR, and that, once transformed to 4S *in vitro*, both agonist-GR and antagonist-GR are able to bind to sequence-specific GRE. Whether other differences between agonist- and antagonist-GR are involved in the

expression of the anti-hormonal activity of RU 486 remains to be elucidated.

Note added in proof: Recent work (J. M. Renoir, J. Devin and E.E.B., in preparation) suggests a similar result when RU 486 is used as an antiprogesterone in the rabbit progesterone system.

Received 26 May; accepted 25 June 1987.

- Willmann, T. & Beato, M. *Nature* **324**, 688–691 (1986).
- Bailly, A., Lepage, C., Rauch, M. & Milgrom, E. *EMBO J.* **5**, 3235–3241 (1986).
- Becker, P. B., Gloss, B., Schmid, W., Strähle, U. & Schütz, G. *Nature* **324**, 686–688 (1986).
- Groyer, A. *et al. Eur. J. Biochem.* **149**, 445–451 (1985).
- Joab, I. *et al. Nature* **308**, 850–853 (1984).
- Catelli, M. G. *et al. EMBO J.* **4**, 3131–3135 (1985).
- Housley, P. R., Sanchez, E. R., Westphal, H. M., Beato, M. & Pratt, W. B. *J. biol. Chem.* **260**, 13810–13817 (1985).
- Sanchez, E. R., Toft, D. O., Schlesinger, M. J. & Pratt, W. B. *J. biol. Chem.*, **260**, 12398–12401 (1985).
- Baulieu, E. E. & Segal, S. J. in *The Antiprogesterin RU 486 and Human Fertility Control* (Plenum, New York, 1985).
- Hager, L. D., McKnight, G. S. & Palmiter, R. D. *J. biol. Chem.* **255**, 7796–7800 (1980).
- Le Bouc, Y. *et al. Endocrinology* **116**, 2384–2392 (1985).
- Schweizer, G., Cadepond-Vincent, F. & Baulieu, E. E. *Biochemistry* **24**, 1742–1749 (1985).
- Mogulewsky, M. & Philibert, D. *J. Steroid Biochem.* **20**, 271–276 (1984).
- Chasserot-Golaz, S. & Beck, G. *J. Steroid Biochem.* **21**, 585–591 (1984).
- Jung-Testas, I. & Baulieu, E. E. *Exp. Cell Res.* **147**, 177–182 (1983).
- Bourgeois, S., Pfahl, M. & Baulieu, E. E. *EMBO J.* **3**, 751–755 (1984).
- Pfahl, M., Sandros, T. & Bourgeois, S. *Molec. cell. Endocr.* **10**, 75–84 (1978).
- Higgins, S. J., Rousseau, G. G., Baxter, J. D. & Tomkins, G. M. *J. biol. Chem.* **248**, 5866–5872 (1973).
- Milgrom, E., Atger, M. & Baulieu, E. E. *Biochemistry* **12**, 5198–5205 (1973).
- Simmons, S. S. Jr & Miller, P. A. *Biochemistry* **23**, 6876–6882 (1984).
- Sablonniere, B., Danze, P. M., Formstecher, P., Lefebvre, P. & Dautrevaux, M. *J. Steroid Biochem.* **21**, 605–614 (1986).
- Green, S. & Chambon, P. *Nature* **324**, 615–617 (1986).
- Marshall, A. J. & Burgoyne, L. A. *Nucleic Acids Res.* **3**, 1101–1110 (1976).
- Albert, B. & Herrick, G. in *Meth. Enzym.* **21**, 198–217 (Academic, London, 1971).

Functional interaction between human T-cell protein CD4 and the major histocompatibility complex HLA-DR antigen

Denise Gay, Paul Maddon*, Rafick Sekaly‡, Mary Anne Talle†, Maurice Godfrey*, Eric Long‡, Gideon Goldstein†, Leonard Chess*, Richard Axel*, John Kappler & Philippa Marrack

Howard Hughes Medical Institute at Denver, National Jewish Center, Denver, Colorado 80206, USA
Departments of Microbiology and Immunology, Biophysics and Genetics, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

* Howard Hughes Medical Institute and Division of Rheumatology, Columbia University College of Physicians and Surgeons, New York, New York 10032, USA

† Ortho Pharmaceuticals Corporation, Raritan, New Jersey 08869, USA

‡ Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892, USA

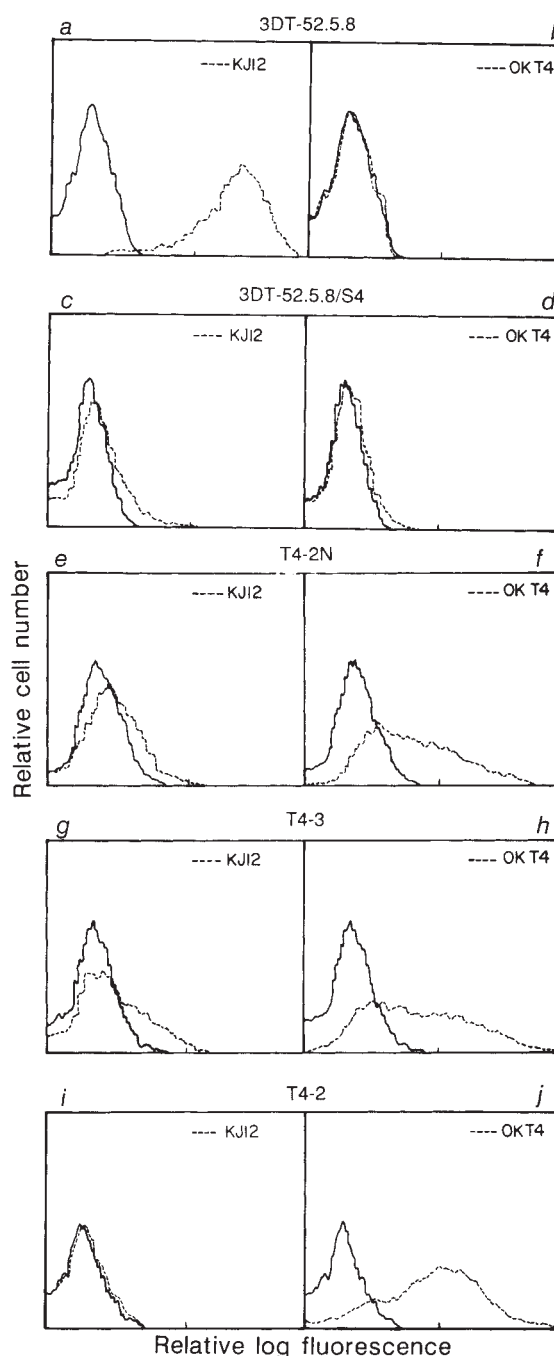
Mature T cells segregate phenotypically into one of two classes: those that express the surface glycoprotein CD4, and those that express the glycoprotein CD8. The CD4 molecule is expressed primarily on helper T cells whereas CD8 is found on cytotoxic and suppressor cells^{1–3}. A more stringent association exists, however, between these T-cell subsets and the major histocompatibility complex (MHC) gene products recognized by their T-cell receptors (TCRs)^{2–4}. CD8⁺ lymphocytes interact with targets expressing class I MHC gene products, whereas CD4⁺ cells interact with class II MHC-bearing targets. To explain this association, it has been proposed that these 'accessory' molecules bind to monomorphic regions of the MHC proteins on the target cell, CD4 to class II and CD8 to class I products^{4–9}. This binding could hold the T cell and its target together, thus improving the probab-

ity of the formation of the trimolecular antigen:MHC:TCR complex. Because the TCR on CD4⁺ cells binds antigen in association with class II MHC, it has been difficult to design experiments to detect the association of CD4 with a class II molecule. To address this issue, we devised a xenogeneic system in which human CD4 complementary DNA was transfected into the murine CD4⁺, CD8⁺ T-cell hybridoma 3DT-52.5.8, the TCR of which recognizes the murine class I molecule H-2D^d. The murine H-2D^d-bearing target cell line, P815, was cotransfected with human class II HLA-DR α , β and invariant chain cDNAs. Co-culture of the parental T-cell and P815 lines, or of one parental and one transfected line resulted in a low baseline response. In contrast, a substantial increase in response was observed when CD4⁺ 3DT-52.5.8 cells were co-cultured with HLA-DR⁺ P815 cells. This result strongly indicates that CD4:HLA-DR binding occurs in this system and that this interaction augments T-cell activation.

A CD4⁺ variant of the murine CD4⁺ T cell hybridoma 3DT-52.5, named 3DT-52.5.8, was obtained by single-cell cloning. This cell line bears high levels of cell-surface receptor as evidenced by staining with KJ12-98 (Fig. 1a), a murine anti-TCR idiotype antibody which identifies the hybrid H-2D^d-restricted receptor unique to 3DT-52.5 and its subclones¹⁰. Recent work by Blackman *et al.*¹¹ has shown that this TCR is a hybrid molecule formed by the fusion partner BW5147 α chain and the parental T cell V β 8-containing β chain. The TCR derived from the incoming T-cell blast can also potentially be expressed by these cells. Therefore, KJ16-133 (ref. 12), an allotypic antibody which recognizes the V β 8 family, was used to detect any surface receptor using the V β 8-containing β chain with a different α chain. Equivalent staining was obtained with both antibodies indicating that the hybrid TCR was the predominant if not the exclusive receptor species present on the 3DT-52.5.8 cell surface (data not shown). The line 3DT-52.5.8/S4, a subclone of 3DT-52.5.8, expressed considerably lower levels of TCR as evidenced by staining with KJ12 and KJ16 (Fig. 1c, data not shown). Neither 3DT-52.5.8 nor 3DT-52.5.8/S4 cells stained with antibodies to either human CD4 (Fig. 1b, d) or murine CD4 (data not shown). It should be noted that numerous murine T-cell hybridomas generated in our laboratory lack surface CD4

Fig. 1 Cytofluorographic analysis of parental and CD4⁺ T-cell hybridomas. T-cell hybridomas were stained with either KJ12-98, a mouse antibody directed against an idiotype determinant of the 3DT-52.5 T-cell receptor (TCR, ref. 10), (1:60), or with OKT4A (ref. 4), a mouse anti-human CD4 monoclonal reagent (1:100). After a 30-min incubation at 20 °C, the cells were washed twice in phosphate buffered saline (PBS) containing 2% fetal bovine serum (FBS), 0.2 M sodium azide and resuspended in fluorescein isothiocyanate (FITC)-conjugated rat anti-mouse immunoglobulin κ light chain antibody (187.1) for another 30 min at 20 °C. After a second set of washes in PBS/FBS/azide, the cells were suspended in 2% paraformaldehyde/0.09 M sodium phosphate/0.2% sodium chloride (pH 7.4). Cells were analysed for fluorescence using a three decade log scale on a Coulter EPICS V cytofluorograph. Control cells stained only with the secondary antibody, FITC-187.1, are represented in all panels as solid lines (—).

Methods. The generation and characterization of the T-cell hybridoma 3DT-52.5 have been described elsewhere¹⁰; 3DT-52.5.8, a CD4⁺ variant of 3DT-52.5, was obtained by single-cell cloning of 3DT-52.5; 3DT-52.5.8/S4 was obtained by single-cell cloning of 3DT-52.5.8 cells presorted for low TCR density. CD4 cDNA isolation, plasmid transfection and infection procedures have been described by Maddon *et al.*¹³.



and CD8 but retain the specificity of their CD4⁺ parents for antigen/MHC. Therefore, CD4 does not appear to contribute to the specificity of the TCR.

The three human CD4⁺ populations (T4-2, T4-2N and T4-3) were obtained by infection of 3DT-52.5.8 cells with murine Maloney leukaemia (Mo-MLV) recombinant virus containing a full-length human CD4 cDNA driven by a murine LTR. This vector also contains the bacterial gene conferring resistance to the antibiotic neomycin¹³. Populations of cells infected with retrovirus were identified by resistance to neomycin and were then analysed for expression of surface CD4 and TCR. As can be seen, these cells expressed reasonable levels of CD4 but considerably lower levels of TCR than the parent 3DT-52.5.8 (Fig. 1e-j). After observing a number of CD4 infectants, some correlation was noted between high CD4 expression and low TCR expression (data not shown) but it is unclear whether this was due to the conditions of Mo-MSV infection or down-regulation of receptor expression by CD4. These infected cells were also phenotypically unstable, losing 80–100% of surface TCR in 5–7 days in culture. As we were unable to stabilize these populations by single-cell cloning, all cell lines were carefully monitored for surface expression of a variety of antigens just before each test of function.

The antigen-presenting cell populations P815/HTR and HLA-DR-transfected P815/HTR (B2-P815) were analysed by cytofluorograph for surface expression of H-2D^d and HLA-DR. As Fig. 2 illustrates, the parental P815 cells and the HLA-DR transfectants exhibited identical staining patterns with respect to cell surface H-2D^d antigen. In contrast, only the HLA-DR transfectants were positive for staining with D1.12, a pan anti-human HLA-DR antibody.

The relative quantity of IL-2 produced by co-culture of the T-cell hybrids with H-2D^d stimulatory cells was used as an indication of relative T-cell activation. As seen in Fig. 3, the mouse CD4⁺ T-cell hybridoma 3DT-52.5.8 produced considerable amounts of IL-2 upon incubation with either the parental P815 or with the HLA-DR⁺ transfectant, B2-P815. Expression of HLA-DR by the P815 transfectants appeared to have no effect upon stimulation of this hybrid. 3DT-52.5.8/S4, the sub-clone of 3DT-52.5.8, produced much lower concentrations of IL-2 when cultured with either P815 population, presumably because this T cell bears much lower levels of cell-surface receptor than its parent. In other respects, however, 3DT-52.5.8/S4 responded similarly to 3DT-52.5.8, responding equally well to both P815 populations.

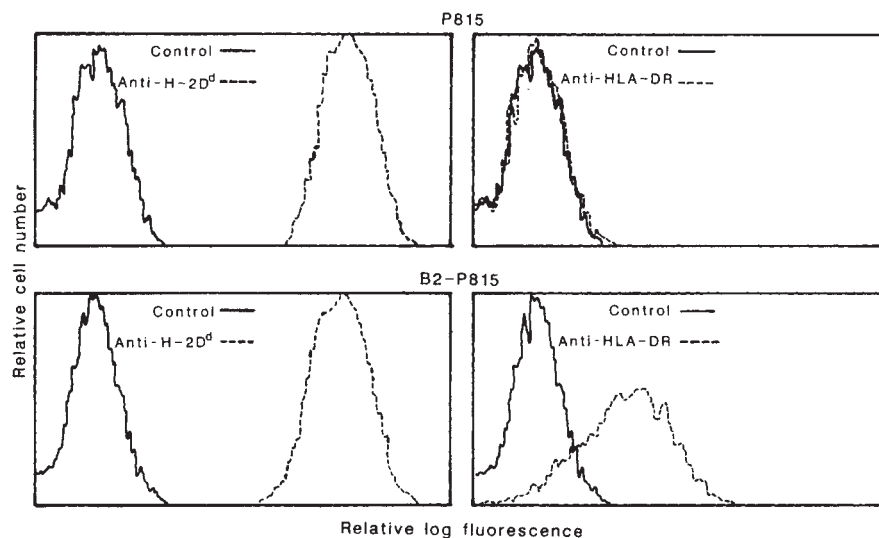
The human CD4⁺ T-cell hybridomas T4-2N and T4-3, both bearing levels of TCR similar to that on 3DT-52.5.8/S4, respon-

ded as poorly as that line to P815 target cells (T4-3 cells with higher TCR density cells did respond slightly better to this stimulus than did either S4 or T4-2N). The T hybrid T4-2, with almost undetectable amounts of TCR, did not produce any IL-2 in response to P815. Expression of CD4 by the CD4-infectants appeared to have no effect upon either TCR recognition of H-2D^d on the P815 cells or subsequent stimulation to produce IL-2.

In contrast, when HLA-DR⁺ P815 cells (B2-P815) were used as the target cells, IL-2 production by T4-2N and T4-3 increased dramatically over that seen with P815 (Fig. 3). More surprisingly, the unresponsive T hybrid, T4-2, now responded significantly. The increase in response could not be explained by an increase in available antigen because (1) both parental and HLA-DR⁺ lines had identical amounts of H-2D^d on their surfaces (Fig. 2) and (2) an increased response was not observed for either CD4⁺ hybrid. The boost in activation required the presence of human

Fig. 2 Cytofluorographic analysis of H-2D^d⁺ P815 populations. The murine DBA/2 mastocytoma line P815/HTR and a human DR1-transfected P815/HTR subpopulation (B2-P815) were stained with either 34-5.8, a mouse anti-murine class I H-2D^d antibody (1:300, ref. 22) or with D1.12, a mouse anti-human HLA-DR antibody (1:100, ref. 21) by the procedure described in Fig. 1. Control cells incubated only with FITC-187.1 are represented as solid lines (—). Both parental P815 cells and HLA-DR⁺ transfectants were 100% positive for 34-5.8 staining and exhibited identical shifts in peak and mean channels.

Methods. Plasmids used for transfection were isolated from a cDNA library constructed with messenger RNA of a human B-cell line as described by Tonnelle *et al.*²³. The library was made in an expression vector carrying the SV40 early promoter and termination sequences. The HLA-DR β chain cDNA corresponds to the HLA-DR1 allelic specificity and has been completely sequenced (Tonnelle). The cDNAs for HLA-DR α and for the invariant chain are full-length and expressible as described by Sekaly *et al.*²⁴. The P815 cell transfections were carried out using calcium phosphate coprecipitation. In a typical experiment, 5×10^6 P815 cells were transfected with 10 μ g supercoiled DNA of each α , β , invariant plasmids and 1 μ g HSVTK plasmid²⁴. Cells were incubated with calcium phosphate reprecipitated DNA in 1 ml for 60 min at 37 °C, then washed and incubated at 37 °C in fresh DMEM/10% FBS. HAT selection was applied 48 h after initiation of the transfection.



CD4 on the T cell as well as human HLA-DR on the stimulating cell.

Antibody inhibition studies illustrated in Table 1 showed that addition of anti-H-2D^d antibody abrogated the response by all the T hybrids to either target population. In contrast, anti-CD4 and anti-HLA-DR antibodies only affected stimulation when the appropriate antigens were on the cell surface and then only when augmented stimulation was noted. Under these conditions, the antibodies usually inhibited down to base-line response. The specificity of the TCR on the CD4⁺ transfectants was not altered to recognize a crossreactive determinant common to murine H-2D^d and human HLA-DR because (1) anti-H-2D^d antibody completely abrogated the response of these cells to HLA-DR⁺ P815 cells and (2) human HLA-DR-1⁺ T-cell blasts could not stimulate the CD4⁺ hybridomas to produce IL-2 (data not shown). Also, the T-cell blast-derived TCR did not contribute to this interaction because 3DT-52.5 negative variants containing only this receptor (KJ16⁺, KJ12⁻) responded to neither P815 nor HLA-DR⁺ P815 cells (data not shown). The above data strongly indicate that a primary function of the CD4 molecule is to bind to the stimulatory cell HLA-DR and physically boost TCR-MHC interaction. Furthermore, the above data substantiates earlier work that MHC restriction of the TCR does not dictate function of CD4⁹. It should be noted, however, that dual recognition of one MHC molecule by both accessory molecule and TCR may be more efficient for T-cell stimulation.

It is increasingly evident that CD4 and CD8 serve a more complicated function in T-cell activation than that of simple adhesion molecules. T-cell stimulation which does not involve the putative target ligands of these molecules can be inhibited and sometimes enhanced with anti-CD4 and anti-CD8 antibodies (refs 14-17 and C. Janeway, in preparation). CD4 cross-linking can induce Ca²⁺ mobilization and antigen-modulation and under certain conditions T cell proliferation¹⁷⁻¹⁹. Also, under certain circumstances, association between the TCR and CD4 can be demonstrated^{19,20}. Additional functions of CD4 and CD8 might therefore be to transduce signals across the T-cell membrane.

Alternatively, if the molecules are in close association with the TCR, concordant recognition of class II or class I MHC molecules by these accessory proteins and the TCR might facilitate formation of the TCR/antigen/MHC tertiary complex. In either case, then, the additional function of the CD4 molecule requires engagement of a ligand. The data presented in this paper supports the view that this ligand is a class II molecule. It is of interest that low TCR expression can be compensated for by presence of CD4 on the T cell and HLA-DR on the stimulatory cell. This indicates that effective interaction between T cells and targets is not exclusively controlled by the affinity and density of TCRs but rather is the result of multiple contributions made by a combination of surface receptors and ligands. This example of the involvement of several different receptor-

Table 1 Units of IL-2 produced by T hybrid cell lines in responses to P815 or HLA-DR⁺ B2-P815 cells

T hybrids:	3DT-52.5.8		3DT-52.5.8/S4		T4-2		T4-2N		T4-3	
Stimulators:	P815	B2-P815	P815	B2-P815	P815	B2-P815	P815	B2-P815	P815	B2-P815
Antibodies to:										
None	640	640	40	40	<10	40	10	80	40	640
H-2D ^d	<10	<10	<10	<10	<10	<10	10	10	10	10
CD4	640	640	40	40	ND	10	10	20	80	320
HLA-DR	640	640	40	40	ND	20	10	10	40	80

Before assaying, T-hybridoma cells were incubated with OKT4B antibody (1:50, ref. 4) and target cells were incubated with either 34-5.8 antibody (1:400) or with cocktail of the following murine anti-human HLA-DR antibodies; BT-2.9 (1:500); D4.22 (1:50); D1.12 (1:200); LG2/72 (1:200); AA3.84 (1:1000) for 1 h at 37 °C²¹. Untreated cells were resuspended in a comparable volume of cell culture medium. IL-2 assay conditions were as described in Fig. 3. Coculture of T hybrid cells and target cells was carried out in the continued presence of the antibodies. These experiments were performed independently from those represented in Fig. 3. Magnitude of IL-2 production by the different T-cell hybrids appears to vary with TCR concentration. ND, not determined.

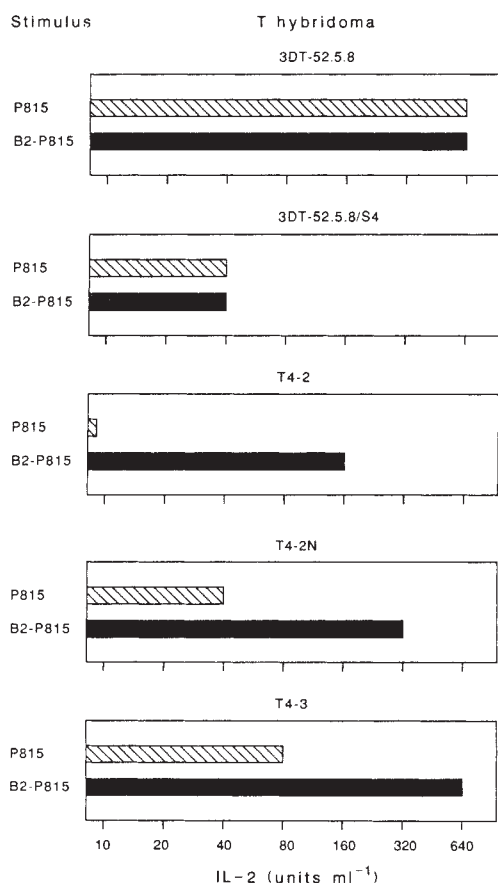


Fig. 3 Response of T-cell hybridomas to HLA-DR⁺ and untransfected P815 cells. T hybrid cells (10^5) were cocultured with 10^5 target cells at 37 °C, 5% CO₂ in 96-well culture plates (Flow); final volume of 300 μ l in culture medium (modified Mishell-Dutton medium/10% FBS/ 5×10^{-5} M 2-mercaptoethanol/50 μ g ml⁻¹ gentamycin (Schering)). After 24 h, the culture supernatants were assayed for the presence of IL-2 by their ability to support the growth of the IL-2-dependent T-cell line HT-2 (ref. 25). The highest of 2-fold serial dilutions capable of maintaining 90% HT-2 cell viability defined the IL-2 concentration of the supernatant. Ten units ml⁻¹ IL-2 was the minimum concentration noted. T-cell hybrids cultured in the absence of the appropriate antigen/MHC stimulus failed to produce any detectable IL-2 (data not shown).

ligand pairs may be a general phenomenon of cell recognition.

We thank Ralph Kubo for helpful suggestions, Judy Pasternak for cytofluorographic assistance, Janice White for technical assistance and John Danilovs for providing human T-cell blasts. This work was supported by the NIH.

Received 12 June; accepted 3 July 1987.

- Reinherz, E., Kung, P., Goldstein, G. & Schlossman, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4061-4065 (1979).
- Golstein, P. *et al. Immun. Rev.* **68**, 5-42 (1982).
- Engleman, E., Benike, C., Glickman, E. & Evans, R. *J. exp. Med.* **153**, 193-198 (1984).
- Biddison, W., Rao, P., Talle, M. A., Goldstein, G. & Shaw, S. *J. exp. Med.* **156**, 1065-1076 (1982).
- Marrack, P. *et al. J. exp. Med.* **158**, 1077-1091 (1983).
- Dialynas, D. *et al. J. Immun.* **131**, 2445-2451 (1983).
- Gay, D., Coeshott, C., Golde, W., Kappler, J. & Marrack, P. *J. Immun.* **136**, 2026-2032 (1986).
- Swain, S., Dutton, R., Schwab, R. & Yamamoto, J. *J. exp. Med.* **157**, 720-729 (1983).
- Greenstein, J., Kappler, J., Marrack, P. & Burakoff, S. *J. exp. Med.* **159**, 1213-1224 (1984).
- Kappler, J., Kubo, R., Haskins, K., White, J. & Marrack, P. *Cell* **34**, 727-737 (1983).
- Blackman, M. *et al. Cell* **47**, 349-357 (1986).
- Haskins, K. *et al. J. exp. Med.* **160**, 452-471 (1984).
- Maddon, P. *et al. Cell* **47**, 333-348 (1986).
- Bekoff, M., Kakiuchi, T. & Grey, H. *J. Immun.* **134**, 1337-1342 (1985).
- Bank, I. & Chess, L. *J. exp. Med.* **162**, 1294-1303 (1985).
- Hunig, T. *J. exp. Med.* **159**, 551-558 (1984).
- Emmrich, F., Strittmatter, U. & Eichman, K. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8298-8301 (1986).
- Ledbetter, J., June, C., Grosmaire, L. & Rabinovitch, P. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1384-1388 (1987).

- Weyland, C., Goronzy, J. & Fathman, G. *J. Immun.* **138**, 1351-1354 (1987).
- Fazekas de St. Groth, B., Gallagher, P. & Miller, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2594-2598 (1986).
- Steel, C. (ed.) *Disease Markers*, Vol. 2 (Wiley, New York, 1984).
- Ozato, K. *et al. Transplantation* **34**, 113-119 (1981).
- Tonnelle, C., De Mars, R. & Long, E. *EMBO J.* **4**, 2839-2857 (1985).
- Sekaly, R., Tonnelle, C., Strubin, M., Mach, B. & Long, E. *J. exp. Med.* (in the press).
- Kappler, J., Skidmore, B., White, J. & Marrack, P. *J. exp. Med.* **153**, 1198-1214 (1981).

Induction of protective immunity against experimental infection with malaria using synthetic peptides

Manuel E. Patarroyo, Pedro Romero, Martha L. Torres, Pedro Clavijo, Alberto Moreno, Alberto Martínez, Raul Rodríguez, Fanny Guzman & Edelmira Cabezas

Instituto de Inmunología, Hospital San Juan de Dios, Universidad Nacional de Colombia, Bogotá, Colombia

Synthetic peptides are potential vaccine candidates because they may be able to induce high antibody titres and specific cellular immune responses against native proteins and thus the whole invading organism¹. In a previous study we showed that immunization with molecules of relative molecular mass (M_r) 155,000 (155K) 83K, 55K and 35K, specific for the late schizont and merozoite stages of *Plasmodium falciparum*, could elicit either partial or total protection in *Aotus trivirgatus* monkeys experimentally infected with *P. falciparum*². Here we have chemically synthesized 18 peptides corresponding to different fragments of these proteins to immunize *Aotus trivirgatus* monkeys. Some peptides gave partial protection from challenge with *P. falciparum* parasites, but none provided complete protection individually. A combination of three partially protective peptides gave complete or almost complete protection, however, suggesting that this particular combination of peptides is a good candidate for a malaria vaccine.

The 155K, 83K, 55K and 35K proteins were isolated from schizont and merozoite lysates, in quantities ranging from 200-400 μ g each, using preparative SDS-PAGE. These proteins showed a high degree of purity as seen by analytical SDS-PAGE and Western blots with hyperimmune sera from malaria patients. The 55K protein provided incomplete protection (defined as a significant delay in the onset of parasitaemia) and the 35K protein partial sterilizing protection (spontaneous control of the experimental infection without drug therapy) on challenge of immunized monkeys with *P. falciparum*².

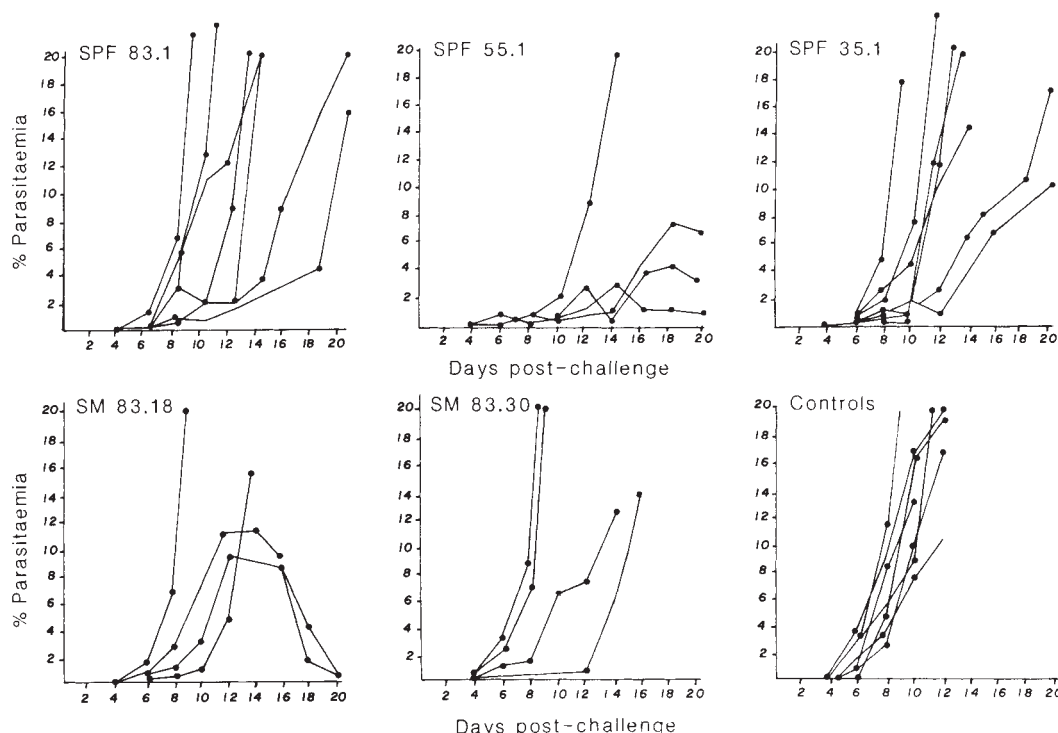
We determined the sequences of the first 21 N-terminal amino acid residues of the 55K and 35K proteins. Based on these data and the amino acid sequences described for the 155K³ or RESA molecule by Coppel *et al.*⁴ together with the complete amino acid sequence of the 195K precursor of the 83K molecule⁵, 18 peptides corresponding to different segments of the 35K, 55K, 83K and 155K proteins were synthesized using Multiple Solid Phase Synthesis⁶. These peptides were chosen randomly to represent different predicted conformations as determined by the Chou and Fasman method⁷ (Table 2), taking into account the suggestion that certain structures can preferentially induce humoral and cellular immune responses^{8,9}.

Groups of four to six Colombian *Aotus trivirgatus* monkeys were immunized with purified peptide (250 μ g) coupled to BSA (250 μ g) on days 0, 30, 45, 60 and 75. Blood samples for antibody studies were taken on days 50, 70 and 80. On day 90, 15 days after the last immunization, the monkeys were challenged with live *P. falciparum*.

Of the 18 different synthetic peptides used for immunization, 15 elicited antibodies against *P. falciparum* schizonts, as detected by immunofluorescence and were found to be immunogenic by

Fig. 1 Development of parasitaemia in monkeys immunized with peptides corresponding to different segments of the 83K, 55K and 35K proteins. Parasitaemia is expressed as the percentage of infected cells. Each line represents a different monkey.

Methods. Groups of 4–6 monkeys were immunized with a single peptide. Peptides (250 µg) had been previously coupled to an equal amount of BSA with glutaraldehyde in a total volume of 250 µl. Conjugated peptides were diluted 1:1 with complete or incomplete Freund's adjuvant (CFA or IFA) and used for subdermal immunization in several parts of the body. Controls received saline solution 1:1 in Freund's adjuvant. Immunizations were carried out on days 0 (with CFA), 30, 45, 60 and 75 (with IFA). The challenge was performed on day 90 with 5×10^6 live fresh *P. falciparum* parasites obtained from another *A. trivirgatus* monkey. This donor was heavily infected with the *P. falciparum* FVO strain (kindly provided by Dr Ruth Nussenzweig) adapted to grow in these monkeys in which it induces a lethal disease. Parasitaemia was monitored daily by examination of smears of fresh peripheral blood treated with the anticoagulant heparin, diluted 1:1 with phosphate buffered saline (PBS), and stained with Giemsa and/or acridine orange.



the dot-blot peptide-antipeptide assay¹⁰. The three peptides that did not elicit antibodies in this immunization scheme, SM 83.15, SM 83.29 and SM 83.23, are sequences from 83K molecule.

Most of the peptides tested failed to provide complete protection against experimental infection regardless of the presence of antibodies (data not shown). Some peptides, however, significantly delayed the onset of the disease in some of the vaccinated animals, suggesting an ability to induce incomplete protection. These peptides were SPf 45.1 and SPf 35.1 (corresponding to the N-terminal sequences of the 55K and 35K molecules

respectively) and SPf 83.1, SPf 83.18 and SPf 83.30 (corresponding to fragments of the 83K protein) (Fig. 1). Some other peptides, for example SPf 83.2, SM 83.23 and SM 83.26 induced incomplete protection in one of four immunized monkeys. Studies are in progress to examine protection in a larger number of animals. None of the peptides used, however, provided complete protection (defined as total absence of parasites in blood smears).

We immunized a new group of monkeys with combinations of two or three of the partially protective synthetic peptides,

Table 1 Antibody titres of immunized monkeys

Sera from monkey no.	Immunized with peptide mixture	IIFA titre	Dot blot-ELISA with		
			SPf 35.1	SPf 55.1	SPf 83.1
229	35.1 + 55.1	20	0	20	ND
255	35.1 + 55.1	80	20	160	ND
287	35.1 + 55.1	160	40	0	ND
251	35.1 + 55.1	160	10	20	ND
275	35.1 + 55.1	40	10	40	ND
288	35.1 + 55.1	5	10	40	ND
289	35.1 + 55.1	5	10	80	ND
286	35.1 + 55.1	160	10	40	ND
295	35.1 + 55.1 + 83.1	10	0	640	640
298	35.1 + 55.1 + 83.1	10	40	320	320
290	35.1 + 55.1 + 83.1	80	0	40	640
291	35.1 + 55.1 + 83.1	80	20	640	640
297	35.1 + 55.1 + 83.1	40	40	160	640
300	35.1 + 55.1 + 83.1	160	160	160	320

Antibody titres of monkeys immunized with the peptide mixtures were determined by indirect immunofluorescence assay (IIFA) and dot-blot ELISA and are expressed as the reciprocal of the serum dilution end-point. In IIFA, air-dried *P. falciparum* erythrocyte cultures with 5% parasitaemia at the schizont stage were used. For the dot-blot ELISA, peptide-antipeptide reactivity was determined using 2 µl of purified peptides at 3 mg ml⁻¹ placed on nitrocellulose sheets and dried. Non-specific reactive sites were neutralized with Tris buffered saline solution (TBS) supplemented with 0.5% Tween 20. Sheets were covered with the appropriate serum dilution for 2 h, washed several times with Tween 20 TBS and incubated with alkaline phosphatase affinity purified anti-human IgG for 2 h. After five more washes with Tween 20 TBS, the reaction was developed with BCIP (5-bromo-4-chlor indolyl phosphate) and NBT (nitroblue tetrazolium), as described by Blake¹². All the pre-immune and control sera were negative by both tests.

Table 2 Synthetic peptides used for immunization

Denomination	Residue number	Original molecule	Sequence
α -Helix			1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
SPf 35.1	N-terminal	35K	Tyr-Gly-Gly-Pro-Ala-Asn-Lys-Lys-Asn-Ala-Gly-OH
SPf 55.1	N-terminal	55K	Asp-Glu-Leu-Glu-Ala-Glu-Thr-Gln-Asn-Val-Tyr-Ala-Ala-NH ₂
SPf 155.1	Repeated octapeptide	155K	Glu-Glu-Asn-Val-Glu-His-Asp-Ala-Tyr-NH ₂
SPf 83.1	43-53	195K	Tyr-Ser-Leu-Phe-Gln-Lys-Glu-Lys-Met-Val-Leu-NH ₂
SM 83.10	27-38	195K	Glu-Leu-Val-Lys-Lys-Phe-Glu-Ala-Leu-Glu-Asp-Ala-OH
SM 83.2	40-52	195K	Leu-Thr-Gly-Tyr-Ser-Leu-Phe-Gln-Lys-Glu-Lys-Leu-Val-OH
SM 83.26	540-552	195K	Leu-Glu-Lys-Leu-Thr-Lys-Ala-Leu-Lys-Tyr-Leu-Glu-Asp-OH
SM 83.28	575-587	195K	Glu-Asn-Glu-Ile-Glu-Thr-Leu-Val-Glu-Asn-Phe-Lys-Lys-OH
SM 83.29	584-595	195K	Asn-Ile-Lys-Lys-Asp-Gly-Glu-Gln-Leu-Phe-Glu-Lys-OH
SM 83.34	731-742	195K	Gln-Ala-Gln-Ala-Gln-Glu-Gln-Lys-Glu-Ala-Gln-Tyr-OH
β -Sheet			
SM 83.31	614-624	195K	Phe-Val-Lys-Val-Gln-Val-Gln-Lys-Val-Leu-Leu-OH
β -Turns			
SM 83.6	78-92	195K	Ser-Gly-Gly-Ser-Val-Ala-Ser-Gly-Gly-Ser-Val-Ala-Ser-Gly-Gly-OH
SPf 83.2	106-117	195K	Tyr-Asn-Ser-Arg-Arg-Thr-Asn-Pro-Ser-Asp-Asn-Gly-OH
SM 83.15	239-250	195K	Glu-Asp-Tyr-Ile-Lys-Lys-Asn-Lys-Lys-Tyr-Ile-Glu-OH
SM 83.23	402-413	195K	Glu-Tyr-Pro-Asn-Gly-Val-Thr-Tyr-Pro-Leu-Ser-Tyr-OH
SM 83.30	595-606	195K	Lys-Lys-Phe-Thr-Lys-Asp-Glu-Asn-Lys-Pro-Asp-Glu-OH
SM 83.33	694-707	195K	Thr-Gln-Gly-Gln-Ser-Asp-Asn-Ser-Glu-Pro-Ser-Thr-Gly-Gly-OH
Random coil			
SM 83.18	277-287	195K	Lys-Leu-Tyr-Gln-Ala-Gln-Tyr-Asp-Leu-Ser-Phe-OH

All the peptides of 195K molecule were synthesized according to the sequence described by Holder *et al.*⁵. The N-terminal sequences of the 35K and 55K molecules were determined by sequencing 250 pm of the purified proteins using a Beckman 890 M sequencer. The residues were identified by HPLC on an ODS column using an acetonitrile/TFA gradient as described previously¹¹. The 18 peptides used for the immunization experiments were synthesized using the multiple solid phase method⁶ in polypropylene bags (74 μ m pore size) with ~150 mg of *p*-methylbenzhydrylamine resin HCl (US Biochemical Corp.). The resin was deprotonated by addition of 5% diisopropylethylamine (DIEA, Aldrich) in methylene chloride (DCM) before introduction of the first amino acid. The coupling cycle was initiated by submerging bags in a solution containing equimolecular amounts of the *t*-boc amino acid (Bachem-Peninsula) and diisopropylcarbodiimide (DIPCDI-Aldrich), in a three-to-fourfold molar excess over available amine in the bag. The reaction was allowed to proceed for 60 min and the product was washed with DCM and isopropanol. The efficiency of the coupling reaction was assessed by picric acid titration. Coupling reactions were repeated if efficiencies fell below 99%. Protective groups of the newly-coupled amino acid were removed with 50% trifluoroacetic acid (TFA-Pierce) in DCM. Reaction products were then washed and amino groups deprotonated with DIEA. Coupling of Asn and Gln was carried out with the addition of 1-hydroxy-benzotriazole hydrate (Aldrich) in dimethylformamide. Amino-acid side groups were protected with Bzl; Arg, His with tosyl, Cys with methoxybenzyl, Lys with Cl-Z, and Tyr with Br-z. The peptides were liberated from the resin by treatment with 2 ml of 10% anisole in anhydrous hydrogen fluoride (HF-Matheson) for 60 min at 0 °C. Twenty-four samples could be cleaved simultaneously. HF was distilled from the reaction and the product was washed 5 times with ethyl ether. The peptides were extracted in 5% acetic acid, analysed and purified by HPLC on an ODS column.

Table 3 Postchallenge parasitaemia in *Aotus* monkeys immunized with synthetic peptides

Monkey no.	% Parasitaemia after challenge on days																												
	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	20	22	24	26	28	30	35	40	45	60	75	90		
Controls																													
358	0.1	0	0.9	0.9	5.3	7.9	10.2	28.4	Q																				
357	0	0	0.1	0	0.5	1.6	2.8	4.0	7.4	19.0	Q																		
359	0.1	0	0.5	0.9	0.9	2.4	2.5	5.4	9.0	9.0	4.5	11.5	12.0	19.0	Q														
Mixture of SPf 35.1 and SPf 83.1																													
229	0	0	0.8	0.5	1.5	5.0	6.6	32.5	Q																				
255	0	0	0.6	1.0	3.7	6.7	10.9	31.0	Q																				
287	0	0	0.8	0.8	1.0	1.6	4.1	6.0	11.8	9.6	9.8	10.0	10.5	Q															
251	0	0	0	0	0	0	0	0.2	0.2	1.2	3.2	ND	11.6	Q															
275	0	0	0	0.1	0.5	0.5	0.5	2.0	4.0	6.5	6.0	6.6	ND	8.3	0.9	0.6	0.5	0.4	0.1	0	0	0	0	0	0	0	0	0	0
288	0	0	0	0	0.2	0.1	0	1.0	0.8	2.3	ND	4.5	5.7	3.1	1.0	1.0	0.5	0.1	0	0	0	0	0	0	0	0	0	0	0
289	0	0	0	0.2	0.1	0.2	1.0	2.3	5.4	10.4	6.8	2.7	5.4	ND	0.2	0.2	0.2	0.1	0	0	0	0	0	0	0	0	0	0	0
286	0	0	0.1	0	0.1	0.4	0.4	3.5	3.7	3.2	ND	0.2	0.2	0	0.2	0.2	0.2	0.2	0.2	0	0	0	0	0	0	0	0	0	0
Mixture of SPf 35.1 and SPf 55.1																													
295	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0.3	4.4	5.5	2.1	0.2	0	0	0	0	0	0	0	0	0
298	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	1.3	4.8	0.9	0.4	0.1	0	0	0	0	0	0	0	0	0	0
290	0	0	0	0.1	0	0	0	0.3	0	0.7	0.1	0.7	0.4	2.5	1.1	0.6	0	0	0	0	0	0	0	0	0	0	0	0	0
291	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
297	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
300	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Postchallenge parasitaemia in *Aotus triuirgatus* monkeys immunized with synthetic peptide mixtures. Parasitaemia is expressed as the percentage of infected cells. Q, initiation of treatment with chloroquine; †, death; ND, not done. The first group of monkeys was immunized with 250 μ g of each peptide, SPf 35.1 and SPf 55.1, conjugated 1:1 to BSA. The second group of monkeys was immunized with a mixture containing equal amounts of the BSA-conjugated peptides SPf 35.1, SPf 55.1 and SPf 83.1. Controls were immunized with saline solution diluted 1:1 in Freund's adjuvant. Immunization and challenge with *P. falciparum* parasites were carried out in the same way as before. Parasitaemia was determined daily.

SPf 55.1 SPf 35.1 and SPf 83.1, using the same immunization schedule as before. Monkeys in this experimental group developed a higher antibody titre compared to the previous group (Table 1).

Four of the eight *Aotus* monkeys immunized with two peptides (55.1 and 35.1) developed the disease similarly to the controls; the remaining four developed parasitaemias lower than 10% and spontaneously recovered. In these animals parasitaemias had not developed by day 180, in contrast to animals immunized with individual peptides.

Three of the six monkeys immunized with the three-peptide mixture (83.1, 55.1 and 35.1) developed very mild infection with parasitaemia maxima of 5%, peaking 10–15 days later than in the control group, followed by spontaneous recovery. The remaining three monkeys of this same group showed no signs of disease and no parasites at all were detected in blood smear samples (Table 3) up to 180 days (shown until day 90) after challenge. There seems to be no direct correlation between antibody titre and protection and, contrary to what has been suggested^{8,9}, there was no correlation of protection with predicted peptide structure. This could be due to the limitations of the method we used to predict these structures.

Recently, Cheung *et al.*¹³ also obtained partial sterilizing protection in a *P. falciparum* vaccine trial in *Saimiri* monkeys with a 31-residue peptide that contains our SPf 83.1 peptide sequence. In contrast to our results, Collins *et al.*¹⁴ described partial sterilizing protection in some *Aotus* monkeys immunized with a genetically engineered fragment of RESA corresponding to a polymeric form of our non-protective SPf 155.1 synthetic monomer. This difference could be due to the fact that we were working with the monomeric form of the repeated octapeptide. The three-peptide mixture may have provided better protection due to the fact that it contains fragments from different blood-stage antigens. One should remember, however, that because of the limited number of monkeys tested with this three-peptide mixture, the possibility remains that the protection offered could be due to a combination of two peptides.

These results show that the combination of the three peptides SPf 35.1 (Tyr-Gly-Gly-Pro-Ala-Asn-Lys-Lys-Asn-Ala-Gly), SPf 55.1 (Asp-Glu-Leu-Glu-Ala-Glu-Thr-Gln-Asn-Val-Tyr-Ala-Ala) and SPf 83.1 (Tyr-Ser-Leu-Phe-Gln-Lys-Glu-Lys-Met-Val-Leu), synthesized according to the amino acid sequences of polypeptides which had already been shown to offer complete or partial immunity against experimental infection, are capable of inducing partial sterilizing or complete immunity in the vaccinated animals. This combination of synthetic peptides is worth considering as a potential vaccine against malaria.

Supported by the Presidency of Colombia. We thank President Belisario Betancur, Diego Pizano Salazar and Dr Mario Calderón Rivera, Dr Armand Hammer and the Occidental Petroleum Company for support, Drs Bruce Merrifield, David Andreu, Richard Houghten and Richard Lerner for teaching us peptide synthesis, Drs Hernan Cuervo, Carlos Parra, Luz Mary Salazar and Clemencia Pinilla for technical assistance, Dr Juan Silva for handling the *Aotus* monkey colony and María Mercedes Zambrano and Fanny Calvo de Simon for help with the manuscript.

Received 28 April; accepted 22 June 1987.

- DiMarchi, R. *et al. Science* **232**, 639–641 (1986).
- Patarroyo, M. E. *et al. Vaccines 87* (eds Brown, F., Chanock, R. & Lerner, R.) (in the press).
- Perlmann, H. K. *et al. J. exp. Med.* **159**, 1686–1704 (1984).
- Coppel, R. L. *et al. Nature* **310**, 789–791 (1984).
- Holder, A. A. *et al. Nature* **317**, 270–273 (1985).
- Houghten, R. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5131–5135 (1985).
- Chou, P. & Fasman, A. *Adv. Enzym.* **47**, 45 (1978).
- Dyson, H. J. *et al. Nature* **318**, 480–483 (1985).
- De Lisi, C. & Berzofsky, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7048–7052 (1985).
- Bittle, J. L. *et al. Nature* **298**, 30–33 (1982).
- Hawke, D., Yuaw, P. & Shively, J. E. *Analyt. Biochem.* **120**, 302–311 (1982).
- Blake, M. S., Johnston, K. H., Russel-Jones, G. J. & Gotschlich, E. C. *Analyt. Biochem.* **163**, 175–179 (1984).
- Cheung, A. *et al. Proc. Natn. Acad. Sci. U.S.A.* **83**, 8328–8332 (1986).
- Collins, W. E. *et al. Nature* **323**, 259–262 (1986).

Induction of *c-fos*-like protein in spinal cord neurons following sensory stimulation

Stephen P. Hunt*, Adrian Pini† & Gerard Evan‡

* MRC Molecular Neurobiology Unit, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

† Department of Pharmacology, University of Cambridge, Cambridge CB2 2QD, UK

‡ Ludwig Institute for Cancer Research, Cambridge CB2 2QH, UK

It has been suggested that the proto-oncogenes *c-fos* and *c-myc* participate in the control of genetic events which lead to the establishment of prolonged functional changes in neurons^{1,2}. Expression of *c-fos* and *c-myc* are among the earliest genetic events induced in cultured fibroblast and phaeochromocytoma cell lines by various stimuli including growth factors, peptides and the intracellular second messengers diacylglycerol, cAMP and Ca²⁺ (refs 3–11). We report here that physiological stimulation of rat primary sensory neurons causes the expression of *c-fos*-protein-like immunoreactivity in nuclei of postsynaptic neurons of the dorsal horn of the spinal cord. Activation of small-diameter cutaneous sensory afferents by noxious heat or chemical stimuli results in the rapid appearance of *c-fos*-protein-like immunoreactivity in the superficial layers of the dorsal horn. However, activation of low-threshold cutaneous afferents results in fewer labelled cells with a different laminar distribution. No *c-fos* induction was seen in the dorsal root ganglia, gracile nucleus or ventral horn. Thus, synaptic transmission may induce rapid changes in gene expression in certain postsynaptic neurons.

Somatosensory information enters the central nervous system (CNS) via primary afferent neurons which synapse on to cells in the dorsal horn of the spinal cord. Activation of small-diameter afferents can lead to the development of long-lasting changes in both the threshold of spinal reflexes^{12,13} and the size of the receptive fields of dorsal horn neurons¹⁴. Because of the suggested link between the expression of the *c-fos* proto-oncogene and the induction of persisting changes in excitable cells^{1,2} we sought evidence that *c-fos* protein is expressed in spinal cord neurons after sensory stimulation.

The following stimuli were applied to activate primary afferent inputs. First, we used cutaneous application and intramuscular injection of the irritant mustard oil (5% in paraffin oil), which produces neurogenic plasma extravasation and stimulation of unmyelinated C polymodal and Aδ myelinated nociceptive afferents^{13,15}. Second, we applied radiant heat (ten localized pulses at 40 °C or 52 °C for 20 s at 90-s intervals) which drives C polymodal nociceptors, although a minor contribution—from heat sensitive Aδ high threshold mechanoreceptors cannot be excluded¹⁶. Third, we used non-noxious brushing of hairs and gentle manipulation of joints (continuous for 15 min) which activates myelinated Aβ and unmyelinated C low-threshold mechanoreceptors, Aβ/δ hairs and Aα proprioceptive afferents^{16–19}.

Two hours after cutaneous application of mustard oil to the hind limb (*n* = 4), p55 *c-fos*-like immunoreactivity was localized to nuclei in a restricted region of the ipsilateral medial dorsal horn of lumbar segments 3 and 4, as is appropriate for this cutaneous input (Fig. 2b). By light microscopy, counterstained cells possessed large nuclei (10 μm maximum diameter) and, at least in deep layers IV and V, were clearly neuronal (Fig. 2e). In more superficial layers cells could not be classified unequivocally as neuronal because of their small size. However, by electron microscopy, all of a sample of 25 superficial cells had dendritic processes and a cytoplasmic and nuclear morphology characteristic of neurons²⁰ (data not shown). Most labelled nuclei were found in layers I (marginal) and II (substantia

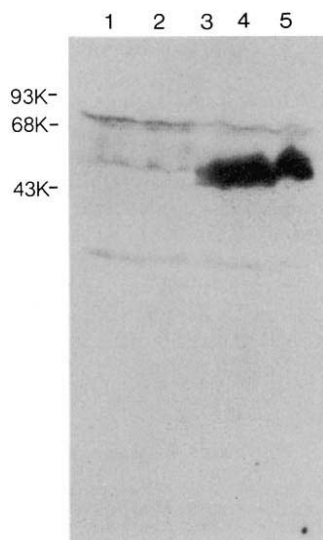


Fig. 1 Characterization of antibody against p55^{c-fos} protein. HeLa S3 cells were growth-arrested by subculture for three days in complete RPMI 1640 medium containing 0.5% fetal calf serum (FCS). The medium was removed and replaced with complete RPMI 1640 medium containing 15% FCS. At various time points equal aliquots of cells were removed and fractionated on an SDS polyacrylamide gel, and compared with an equal number of log-phase cells (track 1). Fractionated proteins were electroblotted onto nitrocellulose paper and the blot probed with antibody raised against the peptide MFSGFNADYEASSRC (denoting amino-acid residues in single letter code) as described below. Tracks are as follows: 1, log-phase; 2, at 0 min; 3, at 30 min; 4, at 60 min; 5, at 120 min. Induction of *c-fos* protein is clearly visible in tracks 3–5. Size markers are, phosphorylase b (93K), bovine serum albumin (68K) and ovalbumin (43K). Antibodies against *c-fos* protein were prepared by immunization with the synthetic peptide Met-Phe-Ser-Gly-Phe-Asn-Ala-Asp-Tyr-Glu-Ala-Ser-Ser-Ser-Arg-Cys. This peptide contains sequences common to the N-terminal region of all known *c-fos* and *v-fos* proteins. The synthetic peptide was conjugated with keyhole limpet haemocyanin using glutaraldehyde and this conjugate was used to immunize New Zealand White rabbits as previously described²⁸. Sera were also assayed for anti-peptide activity by an enzyme-linked immunoadsorbance assay (ELISA)²⁹.

gelatinosa) where the majority of unmyelinated nociceptive afferents terminate^{21,22}, but a scattering of p55^{c-fos}-like-positive neurons was also observed more deeply in layer V, occasionally in layers VI and VII and around the central canal (layer X) (Fig. 2b). No ventral horn nuclei were labelled although some *c-fos* positive neurons (<5% of total) were seen contralaterally. Only occasional labelled nuclei (<5 per section) were found in vehicle treated controls (Fig. 3c) ($n=4$). Immunoreactivity was detectable as early as 20 min ($n=4$) from the start of stimulation but was only just visible at 24 h ($n=4$).

Injection of 10 μ l 5% mustard oil into the medial gastrocnemius muscle ($n=4$) evoked p55^{c-fos}-like immunoreactivity in nuclei largely in the marginal layer I and layers IV–VII of the 4th and 5th lumbar segments of the dorsal horn (Fig. 3c); few labelled nuclei were seen in layer II. Again, this expression is consistent with the somatotopic representation of this muscle in the dorsal horn and with the major terminations of small-diameter nociceptive afferents of muscle in layer I rather than layer II (refs 21–23). No effects were seen in vehicle-injected controls ($n=4$).

Noxious radiant heat localized to the plantar aspect of the hind foot ($n=4$) induced p55^{c-fos}-like immunoreactivity predominantly in layers I and II and layers IV/V in the medial portion of the appropriate lumbar dorsal horn segment^{21,22} and around the central canal (layer X) (Fig. 2a). As previously with mustard oil, only few immunoreactive neurons were seen contra-

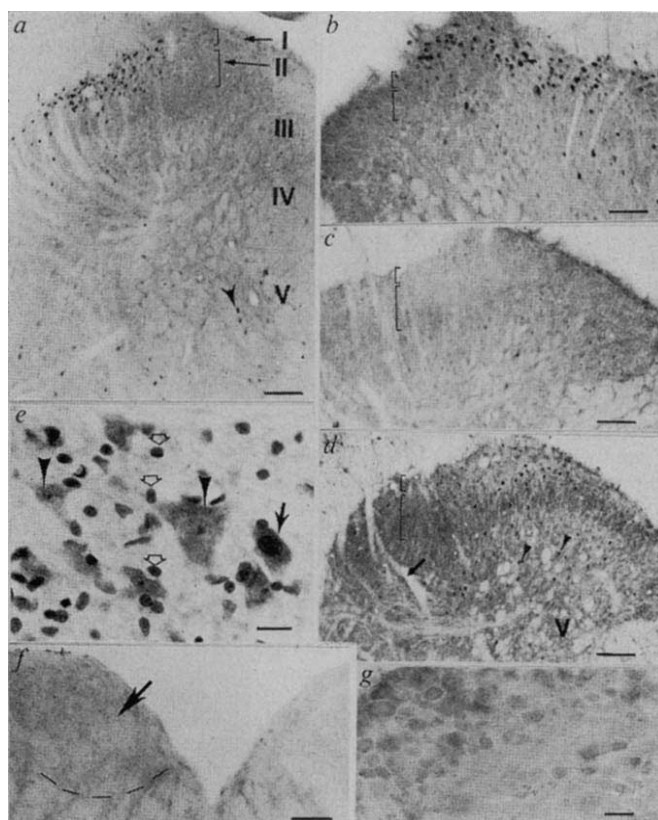


Fig. 2 Immunoperoxidase-stained sections showing p55^{c-fos}-like immunoreactivity in the lumbar dorsal horn of the spinal cord (a–e) following sensory stimulation. a, Radiant heat application (10 times) to the hind digits (52 °C, 20 s on, 90 s off). The layers of the spinal cord are numbered I–V and layers I and II indicated throughout by small and large brackets respectively. Labelled neurons are found in the medial portion of the superficial layers (I and II) and more sparsely in layer V (arrowhead) at 2 h. b, The p55^{c-fos}-like-positive neurons in the medial dorsal horn following application of 5% mustard oil to the hind limb at 2 h. c, Vehicle-treated control. d, Low threshold stimulation of the hind limb, 15 min, survival 2 h. Arrow, incoming myelinated primary afferent fibres. The p55^{c-fos}-like-positive neurons are abundant in layers II–IV (arrowheads) but not in layers I or V. e, High-power photomicrograph of a cresyl violet counterstained section. Filled arrow, the only p55^{c-fos}-like positive neuron. Unlabelled cells are indicated by the black arrowheads (neurons) and open arrows (glia) respectively. f, g, Lack of p55^{c-fos}-like immunoreactivity in the nucleus gracilis (arrow in f) and L4 dorsal root ganglion following electrical stimulation (at 1 Hz) of the ipsilateral sciatic nerve at C fibre strength for 10 min. Scale bars: a, 80 μ m; b, 60 μ m; c, d, 70 μ m; e, 15 μ m; f, 80 μ m; g, 40 μ m. A second antibody raised against a synthetic peptide corresponding to residues 245–253 of human *c-fos* protein gave comparable results (data not shown). Methods as in Fig. 3.

laterally whereas radiant heat at 40 °C was completely ineffective ($n=2$). Following non-noxious brushing and manipulation of the hind limb ($n=4$) labelled nuclei were seen mainly in layers II to IV and rarely in layer I (Figs 2d and 3b). No *c-fos*-like immunoreactivity was observed in the dorsal root ganglion, nucleus gracilis or ventral horn following noxious chemical or radiant heat stimulation.

Observations of the effects of electrical stimulation of nerve trunks were compromised by the finding of massive *c-fos*-like induction due to surgery in sham stimulated controls ($n=4$). Further experiments ($n=4$) revealed that a clean 2-cm cutaneous incision was sufficient to induce extensive *c-fos*-like immunoreactivity in layers I, II, V, and X. However, electrical stimulation¹⁶ of the sciatic or saphenous nerves which recruited

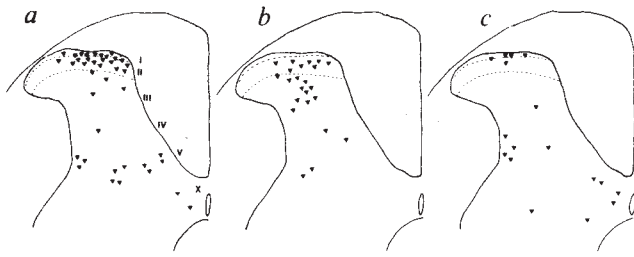


Fig. 3 The stimulus-dependent laminar distribution of p55^{c-fos}-like-positive neurons. Sections are drawn from the most extensively labelled regions of the lumbar spinal cord. *a*, Radiant heat pulses (52°C) to the plantar surface of the foot; *b*, low-threshold stimulation of the hind limb; *c*, intramuscular injection of 10 µl of 5% mustard oil. Each triangle represents three labelled neurons.

Methods. Sprague-Dawley rats (180–250g) were deeply anaesthetized with Equithesin (4 ml per kg) intraperitoneally (i.p.) before primary afferent stimulation and perfusion. At 20 min to 24 h, but most usually 2 h after stimulation, animals were perfused intracardially with 100 ml sodium phosphate buffered saline (0.1 M, pH 7.4) followed by 200–300 ml of 4% paraformaldehyde in phosphate buffer (0.1 M, pH 7.4, PB). Tissue was removed, postfixed in the latter solution for 2 h and washed overnight in 30% sucrose in PB. Frozen sections were incubated overnight in rabbit antisera diluted 1:1,500 in 0.1 M Tris-saline pH 7.4 (TBS) and 0.3% Triton X-100. Tissue was washed in TBS then incubated with sheep antirabbit antiserum (Vector) diluted 1:200 in TBS plus 0.3% Triton X-100 for 60 min, rewashed in TBS and finally incubated with avidin-biotin-horse-radish peroxidase complex (Vector) for 1 h in TBS/0.3% Triton X-100. Sections were washed for 1 h in TBS and incubated in diaminobenzidine (100 µg per ml) with 0.01% hydrogen peroxide for 5 min before being washed in PB, and mounted on glass slides. Some tissue was prepared for electron microscopy as previously described³⁰. Control tissue was prepared with normal rabbit serum and primary antisera adsorbed against 10 µg ml⁻¹ of native peptide both of which abolished all nuclear staining.

all unmyelinated and/or all myelinated axons (determined from compound action potentials) did not induce p55^{c-fos}-like immunoreactivity in the nucleus gracilis (Fig. 2f) or ventral horn (which receive powerful monosynaptic inputs from Aαβ afferents)^{17–19} or dorsal root ganglia (Fig. 2g). It therefore appears that depolarisation is not sufficient for c-fos induction in dorsal root ganglia, nucleus gracilis and ventral horn and that the expression of c-fos protein is dependent upon either the neurochemistry of synaptic transmission or upon the postsynaptic propensity for induction or both.

We cannot distinguish precisely between the contribution to expression of c-fos protein signalled by unmyelinated versus myelinated inputs or whether expression is monosynaptic or polysynaptic. However, by far the most extensive labelling is generated by noxious stimuli in those regions of the dorsal horn which receive small-diameter primary afferents. The much weaker activity induced in layers II to IV by non-noxious stimuli is consistent with actions mediated from the terminals of low-threshold Aδ and C primary afferents^{18,24}.

Our observations clearly show that in the dorsal horn, primary afferent input generates rapid trans-synaptic expression of c-fos-like immunoreactivity whose laminar distribution is related to the nature of the sensory stimulus. We suggest that c-fos protein is induced by neurotransmitters, possibly peptides, known to be present in small-diameter afferents²⁵ which are the major inputs to those regions of greatest p55^{c-fos}-like immunoreactivity in the dorsal horn. That low-threshold inputs could induce c-fos protein through peptidergic interneurons²⁵ appears likely because many cells of the dorsal horn and dorsal column nuclei which respond only to non-noxious inputs are readily excited by ionophoretic application of neuropeptides^{26,27}.

Thus, sensory transmission causes rapid induction of the c-fos proto-oncogene in the mammalian CNS. This gene is thought

to function in the control of both mitogenesis and differentiation and is known to be induced by the intracellular second messengers diacylglycerol^{3,5} cyclic AMP (refs 5 and 9) and Ca²⁺ (ref. 7) in response to external signals such as peptides¹⁰ and growth factors^{4,6}. Although the functioning of c-fos is poorly understood, our demonstration of its synaptic induction is the first evidence that sensory information can influence gene expression in neurons of the adult CNS.

We thank Annette Bond for technical assistance, Dr B. Verrier and Dr R. Muller for antisera used in pilot studies, Mrs Gerry Carter for typing the manuscript and Professor Eric Barnard, Dr Stephen Davies and Dr Graeme Henderson for comments on the manuscript. A.P. thanks Dr Ron Ryall for the use of his facilities. A.P. is an MRC Training Fellow.

Received 21 April; accepted 17 June 1987.

- Goel, P., Castellucci, V. F., Schacher, S. & Kandel, E. R. *Nature* **322**, 419–422 (1986).
- Berridge, M. *Nature* **323**, 294–295 (1986).
- Greenberg, M. E. & Ziff, E. B. *Nature* **311**, 433–437 (1984).
- Greenberg, M. E., Greene, L. A. & Ziff, E. B. *J. biol. Chem.* **260**, 14101–14110 (1985).
- Krujjer, W., Schubert, D. & Verma, I. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7330–7334 (1985).
- Milbrandt, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4789–4793 (1986).
- Morgan, J. I. & Curran, T. *Nature* **322**, 552–555 (1986).
- Muller, R., Bravo, R., Burckhardt, J. & Curran, T. *Nature* **312**, 716–720 (1984).
- Ran, W., Dean, M., Levine, R. A., Henkle, C. & Campisi, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8216–8220 (1986).
- Rozengurt, E. *Science* **234**, 161–166 (1986).
- Stumpo, D. J. & Blackshear, P. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9453–9457 (1986).
- Woolf, C. J. *Nature* **306**, 686–688 (1983).
- Woolf, C. J. & Wall, P. D. *J. Neurosci.* **6**, 1433–1442 (1986).
- Cook, A. J., Woolf, C. J., Wall, P. D. & McMahon, S. B. *Nature* **325**, 151–153 (1987).
- Heapy, C. G., Jamieson, A. & Russell, N. J. W. *Br. J. Pharmac.* **90**, 164P (1987).
- Lynn, B. & Carpenter, S. E. *Brain Res.* **238**, 29–43 (1982).
- Brown, A. G. *Organisation in the Spinal Cord* (Springer, Heidelberg, 1981).
- Brown, A. G. in *Spinal Cord Sensation* (eds Brown, A. G. & Rethelyi, M.) 23–32 (Scottish Academic Press, Edinburgh, 1981).
- Lynn, B. in *Biochemistry and Physiology of the Skin* (ed. Goldsmith, L. A.) 654–684 (Oxford University Press, New York, 1983).
- Peters, A., Palay, S. L. & Webster, H. deF. *The Fine Structure of the Nervous System* (Saunders, Philadelphia, 1976).
- Molander, C. & Grant, G. *Neuroscience* **19**, 297–312 (1986).
- Swett, J. & Woolf, C. J. *Comp. Neurol.* **231**, 66–77 (1985).
- Craig, A. O. & Mense, S. *Neurosci. Lett.* **41**, 233–238 (1983).
- Sugiura, Y., Lee, C. L. & Perl, E. R. *Science* **234**, 358–361 (1986).
- Hunt, S. P. in *Chemical Neuroanatomy* (ed. Emson, P. C.) 53–84 (Raven, New York, 1983).
- Pini, A. & Ryall, R. W. *J. Physiol., Lond.* **378**, 42P (1986).
- Krnjevic, K. in *Substance P* (eds von Euler, U. S. & Pernow, B.) 217–239 (Raven, New York, 1977).
- Ramsey, G., Evan, G. I. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7742–7746 (1984).
- Evan, G. I., Lewis, G. K., Ramsey, G. & Bishop, J. M. *Molec. cell. Biol.* **5**, 33610–33616 (1985).
- Ohara, P. T., Lieberman, A. R., Hunt, S. P. & Wu, J.-Y. *Neuroscience* **8**, 189–211 (1983).

An isoprenaline activated sodium-dependent inward current in ventricular myocytes

T. M. Egan*, D. Noble, S. J. Noble,
T. Powell & V. W. Twist

University Laboratory of Physiology, Parks Road,
Oxford OX1 3PT, UK

In the heart, catecholamines affect pacemaker activity by shifting the activation curve for the nonspecific inward current^{1–4} and increasing both the calcium current^{5–8}, and the delayed potassium current⁹. We report here that in mammalian ventricle there is another mechanism that seems to involve a sodium-dependent inward current. This is elicited by agents that increase intracellular cyclic AMP concentration, such as the β-adrenergic agonist isoprenaline, and is unaffected by agents which block the three currents listed above, but is absent when external sodium is replaced with tetramethylammonium. Most interestingly, the intracellular pathway(s) linking the β-receptor(s) to activation of the Ca current and the Na-dependent current, which in both cases

* Present address: Max-Planck-Institut für Psychiatrie, Abteilung Neurophysiologie, Klopferstr. 18 A, 8033 Planegg-Martinsried, FRG.

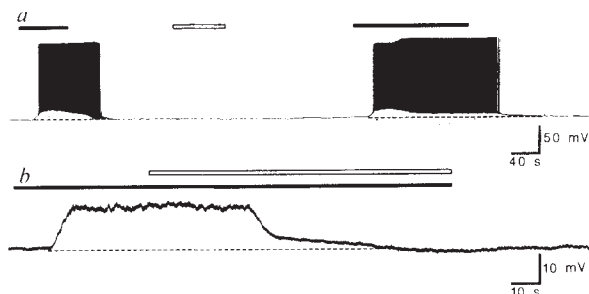


Fig. 1 Isoprenaline depolarizes ventricular myocytes through an action on β -adrenergic receptors. *a*, Superfusion of (-)isoprenaline (solid bar) (100 nM) depolarized the cell (resting membrane potential -82 mV) thereby initiating repetitive action potential discharge. Superfusion of (+)isoprenaline (open bar) was without effect. *b*, Superfusion of 100 nM (-)isoprenaline (solid bar) in this cell (resting membrane potential -80 mV) caused a 15-mV depolarization which did not result in repetitive firing. This effect of isoprenaline was blocked upon addition of the β -adrenoceptor antagonist propranolol (100 nM, open bar). Not shown is the return of isoprenaline-induced depolarization after 20 min wash of the propranolol.

Methods. Experiments were performed on single ventricular myocytes isolated from adult guinea-pig hearts by enzymatic perfusion^{22,23}. The cells were studied *in vitro* in a flowing (1–2 ml min⁻¹) Krebs-Ringer solution heated to a temperature of 34–36 °C. The composition of the Krebs solution was (except where noted) (in mM): NaCl 144, KCl 5.4, NaH₂PO₄ 0.3, MgCl₂ 1, CaCl₂ 2.5, glucose 5.6, HEPES 5, buffered to pH 7.4 with NaOH. Membrane potential and current were recorded using a single conventional microelectrode filled with either KCl (1–2 M), K₂SO₄ (0.5 M), CsCl (1 M) or a KCl (1 M)/EGTA (100 mM) mixture, and a switch-clamp amplifier (Axoclamp-2, Axon Instruments). Microelectrode resistances measured 8–12 M Ω when 2 M KCl was used as the filling solution. Drugs were applied by addition of known concentrations to the superfusate. In some experiments isoprenaline was applied by pressure ejection (10–100 ms of 70–140 kPa pressure) of a very small amount (a few nanoliters) of an isoprenaline-Ringers mixture (pipette concentration 1 μ M) from a micropipette (tip diameter 10–15 μ m) in the superfusate above the experimental cell. Cs and EGTA were applied intracellularly by diffusion from the recording electrode. The effects reported here were seen in >90% of the cells included in this study ($n=171$) and in all the cells in which the well-documented increase in i_{Ca} was also seen. Unless otherwise noted, each experiment was repeated at least five times.

presumably involves the intracellular concentration of cAMP, differ, as isoprenaline causes a persistent augmentation of the calcium current whereas the Na-dependent current often fades. These effects of isoprenaline are antagonized by acetylcholine. In unclamped cells, the Na-dependent current depolarizes the membrane to the potential range at which repetitive firing occurs. It may therefore be involved in the generation of ventricular arrhythmias.

Figure 1a shows the effect of isoprenaline (10–1,000 nM). Superfusion of the biologically active (-) stereoisomer of isoprenaline (100 nM) caused a depolarization large enough to initiate repetitive activity, whereas the inactive (+) isomer (100 nM) was without effect. This action of (-) isoprenaline was blocked by the β -adrenoceptor antagonist propranolol (100 nM, Fig. 1b).

Under voltage-clamp isoprenaline (10–1,000 nM) generated a dose-dependent inward current (0.1–1.0 nA) having the same time course as the drug-induced membrane depolarization (Fig. 2a). Both membrane conductance and noise increased. The current amplitude became smaller at potentials positive to rest (Fig. 2b–e). Unfortunately, the fact that isoprenaline affects other currents (the Ca current, i_{Ca} and the delayed K current, i_K) made reliable determination of the reversal potential for the

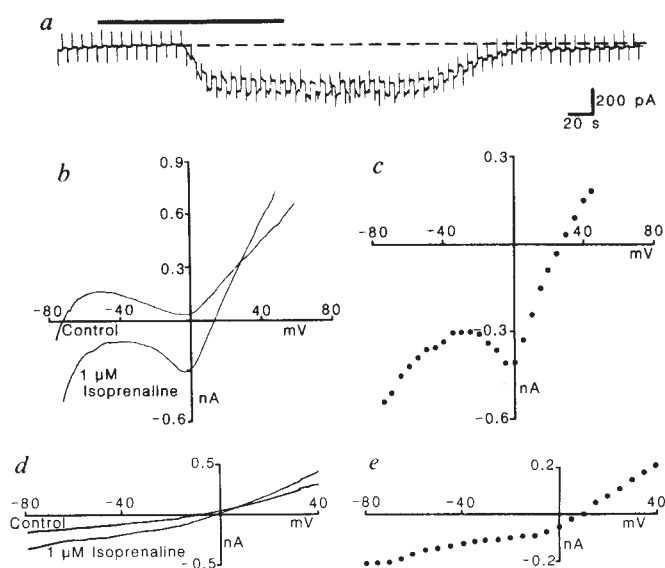


Fig. 2 Isoprenaline-induced inward current under voltage-clamp. *a*, Superfusion of isoprenaline (1 μ M, solid bar) evoked an inward current and an increase in membrane conductance. The external solution was nominally Ca-free and contained elevated [Mg]_o (10 mM) and Mn (1 mM). This allowed the measurement of conductance changes underlying the drug-induced current which were largely uncontaminated by the effect of isoprenaline on calcium conductance. To measure conductance, the cell was stepped for 3 s from a holding potential of -40 mV to -30 mV once every 6 s. In the absence of protocols used to block i_{K1} , we found it easiest to measure drug-induced changes in conductance at potentials (such as -40 to -30 mV) where i_{K1} contributes little to the steady-state membrane current-voltage curve. *b*, Current-voltage ($I-V$) relationship for a single myocyte under voltage-clamp before and during superfusion of isoprenaline (1 μ M). Membrane current was measured during a voltage-ramp from the resting potential to about +50 mV applied at the rate of 20 mV s⁻¹. *c*, Net current-voltage curve evoked by isoprenaline in the experiment in *b*. This curve was obtained by subtracting the $I-V$ curve during drug application from that obtained before application. *d*, Current-voltage relationship before and during superfusion of isoprenaline (1 μ M) obtained from a cell in which most of the i_{Ca} and i_K were blocked by removal or addition of the following ions and drugs to the external superfusate (in mM): 0 Ca²⁺, 2 Ba²⁺, 2 Mg²⁺, 1 Cd²⁺, 0.010 nifedipine, 5 Cs⁺. The recording electrode contained 1 M CsCl. *e*, Net current-voltage curve for the current caused by isoprenaline in the experiment in *d*.

inward current difficult. This problem is clear from the experiment of Fig. 2b and c where the total membrane current over a range of potentials was measured before and during application of isoprenaline (1 μ M). Addition of isoprenaline had three clear effects. First, at potentials more negative than +27 mV, isoprenaline caused a net inward current; at more positive potentials, the net current was outwards. Second, isoprenaline increased the amplitude of a non-inactivating component of i_{Ca} . Third, isoprenaline increased i_K , producing an outward rectification at positive potentials. Thus the measured reversal potential for the current reflects these multiple effects of isoprenaline. Even in solutions which block a large proportion of i_{Ca} and i_K (Fig. 2d and e) but leave the inward current unaffected (see below), we could not say with confidence that we measure a true reversal potential for the inward current uncontaminated by other currents. Although this regime did linearize the drug-induced current-voltage relationship at negative potentials (-80 mV to about -10 mV), it did not block a significant outward rectification at potentials positive to -10 mV. We are unsure whether this outward rectification reflects the true nature of the drug-induced current or whether it results from residual i_K .

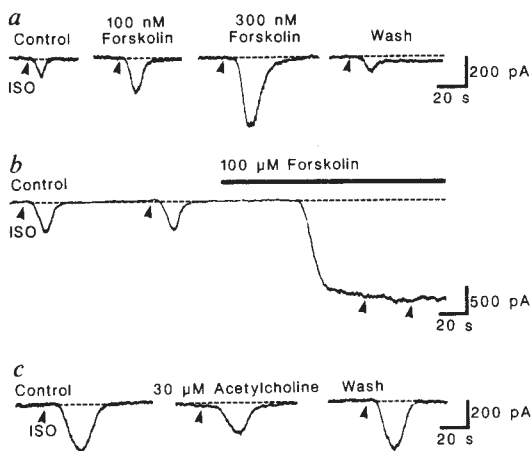


Fig. 3 The effect of forskolin and acetylcholine on the Na-dependent inward current. *a*, A submaximal inward current was evoked by ejection of isoprenaline (arrowheads) from a pressure-loaded micropipette. Superfusion of 100 nM and 300 nM forskolin caused no Na-dependent inward current on their own but augmented that caused by isoprenaline. Maximal augmentation was seen after about 5 min exposure to forskolin. The isoprenaline evoked current returned to its control amplitude after 2 min wash of the forskolin. *b*, Same cell as in *a*. Superfusion of a large concentration of forskolin (100 μ M) occludes the inward current evoked by pressure ejection of isoprenaline. *c*, Superfusion of acetylcholine (30 μ M) caused no change in holding current but reduced the Na-dependent inward current evoked by pressure application of isoprenaline.

In sheep cardiac Purkinje fibres, isoprenaline causes a membrane depolarization by shifting the activation curve of the non-specific inward current i_f (refs 2 and 4); these effects are reduced by external Cs^+ . The isoprenaline-induced current in our ventricular cells was not affected by Cs (5 mM). It was also unaffected by the K-channel blockers tetraethylammonium (20 mM), Ba^{2+} (2 mM) and internal Cs^+ . Likewise, the inward current was not blocked by the calcium channel blockers Cd^{2+} (100 μM –1 mM), Mg^{2+} (5–10 mM) and nifedipine (10–20 μM), or by elimination of Ca^{2+} from the external medium; it was also unaffected when the concentration of intracellular Ca^{2+} , $[\text{Ca}^{2+}]_i$, was buffered with an amount of EGTA sufficient to abolish contraction. Addition of strophanthidin (10 μM) or elimination of K^+ from the external medium, both of which inhibit the Na^+/K^+ pump, were without effect, as was substitution of isethionate for extracellular chloride ($n=3$). However, the inward current with its associated increase in membrane conductance was dependent on the concentration of extracellular Na^+ ($[\text{Na}^+]_o$) being absent when external Na^+ was replaced by an equimolar concentration of the impermeant ion tetramethylammonium. These results suggest that the inward current is generated by mechanisms distinct from the well-documented effects of catecholamines on i_f , i_{Ca} , i_{K} and the Na^+/K^+ pump^{3–10} and that Na is the major charge carrier.

Isoprenaline increases i_{Ca} by activating adenylate cyclase and increasing $[\text{cAMP}]_i$ (refs 7, 8, 10, 11). Our evidence suggests that $[\text{cAMP}]_i$ is also involved in generation of the Na-dependent inward current described here, because any agent which increases $[\text{cAMP}]_i$ either directly by stimulating adenylate cyclase (for example forskolin, at 600 nM–100 μM ; see Fig. 3*b*) or indirectly by inhibiting phosphodiesterase (isobutylmethylxanthine, 0.020–1 mM) activated the Na^+ -dependent current. In addition, lower concentrations of forskolin (<600 nM; Fig. 3*a*) and IBMX (<20 μM), which had no effect of their own, potentiated the action of isoprenaline. This is strong evidence for $[\text{cAMP}]_i$ mediation¹².

The effects of isoprenaline and forskolin were non-additive

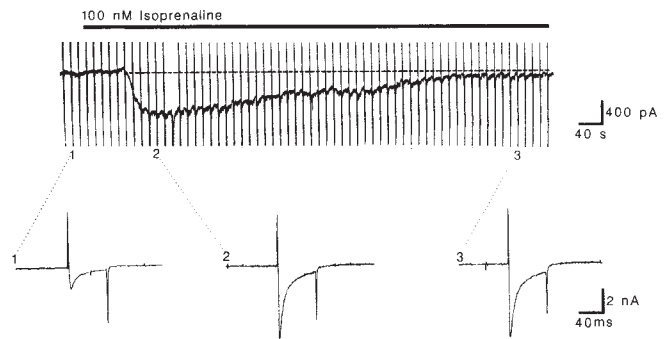


Fig. 4 Separation of the effect of isoprenaline on the Na-dependent inward current and i_{Ca} . Membrane current was measured in a single ventricular myocyte held under voltage-clamp at its resting potential of -75 mV: i_{Ca} was evoked once every 10 s by voltage-clamp command steps, first to -40 mV for 600 ms to inactivate fast i_{Na} and then to 0 mV for 50 ms to evoke i_{Ca} . Dotted line, zero holding current. Superfusion of isoprenaline evokes both the flow of inward current and augmentation of i_{Ca} . However, whereas the Na-dependent inward current slowly fades during superfusion of the agonist, the augmentation of i_{Ca} remains.

(Fig. 3*b*). Thus, a very high concentration of forskolin (100 μM) caused a large inward current which occluded the inward current which would have been caused by pressure ejection of isoprenaline. Again, this result suggests a common mechanism underlying the currents evoked by isoprenaline and forskolin.

An additional feature of the inward current caused by either forskolin or isoprenaline was that it was reduced 30–50% by acetylcholine (10–30 μM), although acetylcholine on its own caused no change in holding current (Fig. 3*c*). The mechanism of this antagonism is unknown but may involve a reduction in $[\text{cAMP}]_i$ by acetylcholine, as is the case when isoprenaline and acetylcholine interact on cardiac i_{Ca} (refs 10, 13 and 14).

Whereas isoprenaline apparently causes both the flow of a Na-dependent inward current and augmentation of i_{Ca} by increasing $[\text{cAMP}]_i$, there is a distinct difference in the respective time courses. Thus, in Fig. 4, the drug-induced inward current slowly faded, until after 8 min of drug-application no Na-dependent inward current remained. This fade occurred in about 50% of the cells exposed to either isoprenaline or forskolin. Desensitized cells were insensitive to subsequent applications of either isoprenaline or forskolin for up to 30 min after the first, desensitizing application. The mechanism(s) underlying the fade in Na-dependent current are unknown, but it seems unlikely that the cause is general depletion of intracellular cAMP or its precursor ATP because: (1) most importantly, i_{Ca} does not fade, and (2) depletion of ATP results in identifiable symptoms in ventricular cells which we fail to detect after our fade^{8,15}.

In conclusion, isoprenaline at physiological concentrations can activate an inward sodium current at the resting potential, thus generating repetitive firing in ventricular myocardium. This finding may well prove to be of clinical importance because adrenergic amines are thought to be involved in the development of arrhythmia and fibrillation following acute coronary occlusion^{16–20}, as evidenced by increased catecholamine levels in the coronary sinus blood. It has, however, been difficult to account for the mechanism underlying the arrhythmia²¹. Our results provide for the first time a direct membrane action by which such arrhythmias could occur. A feature of more speculative importance is that this current mechanism could also be involved in natural pacemaker activity. The relative contribution of a background inward current is greater in sinus node tissue, which is in part attributable to the near absence of the inward rectifier current i_{K1} in this tissue, but may also be attributable to the maintained presence of a current such as we describe here. Further experiments are required to clarify this question,

and these will be of particular importance because the background inward current in the heart is still very poorly understood.

Supported by grants from the British Heart Foundation, the MRC, the NIH and the Wellcome Trust.

Received 30 March; accepted 26 June 1987.

- Hauswirth, O., Noble, D. & Tsien, R. W. *Science* **162**, 916-917 (1968).
- DiFrancesco, D. *Prog. Biophys. molec. Biol.* **46**, 163-183 (1985).
- Terris, S., Wasserstrom, J. A. & Fozzard, H. A. *Am. J. Physiol.* **251**, H1056-H1061 (1986).
- Glitsch, H. G. & Rasch, R. *Pflügers Arch. ges. Physiol.* **406**, 144-150 (1986).
- Reuter, H. *Nature* **301**, 569-574 (1983).
- Tsein, R. W. *et al. J. molec. cell. Cardiol.* **18**, 691-710 (1986).
- Kameyama, M., Hofmann, F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **405**, 285-293 (1985).
- Kameyama, M., Hescheler, J., Hofmann, F. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **407**, 123-128 (1986).
- Brown, H. F. & Noble, S. J. *J. Physiol., Lond.* **238**, 51-53P (1974).
- Fischmeister, R. & Hartzell, H. C. *J. Physiol., Lond.* **376**, 183-202 (1986).
- Irisawa, H. & Kokubun, S. *J. Physiol., Lond.* **338**, 321-337 (1983).
- Seamon, K. B. & Daly, J. W. *Trends Pharmac. Sci.* **4**, 120-123 (1983).
- Watanabe, A. M. & Besch, H. R. Jr *Circulation Res.* **37**, 309-317 (1975).
- Löffelholz, K. & Pappano, A. J. *Pharmac. Rev.* **37**, 1-24 (1985).
- Noma, A. *Nature* **305**, 147-148 (1983).
- Richardson, J. A., Woods, E. F. & Bagwell, E. E. *Am. J. Cardiol.* **5**, 613 (1960).
- Russell, R. A., Crafoord, J. & Harris, A. S. *Am. J. Physiol.* **200**, 995-998 (1961).
- Mathes, P., Cowan, C. & Gudbjarnason, S. *Am. Heart J.* **81**, 211-216 (1971).
- Mathes, P. & Gudbjarnason, S. *Am. J. Physiol.* **220**, 27-32 (1971).
- Staszewska-Barczak, J. & Ceremuzynski, L. *Clin. Sci.* **34**, 531-539 (1968).
- Angelakos, T. E. & Torchiana, M. L. in *Cardiac Arrhythmias* (eds Dreifus, L. S. & Likoff, W.) 505-516 (Grune and Stratton, New York, 1973).
- Mitchell, M. R., Powell, T., Terrar, D. A. & Twist, V. W. *Br. J. Pharmac.* **81**, 543-550 (1984).
- Powell, T., Terrar, D. A. & Twist, V. W. *J. Physiol., Lond.* **302**, 131-153 (1980).

Prostaglandins and the synergism between VIP and noradrenaline in the cerebral cortex

Nicolas C. Schaad*, Michel Schorderet*†
& Pierre J. Magistretti*†

* Département de Pharmacologie, Centre Médical Universitaire,
1211 Geneva, Switzerland

† Ecole de Pharmacie, 1000 Lausanne, Switzerland

We have previously shown that vasoactive intestinal peptide (VIP) and noradrenaline (NA) interact synergistically to increase cyclic AMP levels in mouse cerebral cortical slices¹. The pharmacological mechanism of this synergism is the potentiation by NA, through α_1 adrenergic receptors, of the stimulatory effect of VIP on cAMP formation². A similar interaction has been confirmed in guinea pig cerebral cortex³ and in discrete nuclei of the rat hypothalamus⁴. Furthermore VIP and NA interact synergistically to depress the spontaneous activity of identified neurons in rat neocortex⁵. At the cellular level, this synergistic interaction suggests that VIP- and NA-containing neuronal systems may converge, at least in part, on the same target cells to increase cAMP levels in the cerebral cortex. At the molecular level, the interaction may occur at various steps in signal transduction, between receptors, intramembrane transduction processes or intracellular effector mechanisms. Here we report that the α_1 -adrenergic potentiation of the increases in cAMP elicited by VIP involves the formation of arachidonic acid metabolites and is mimicked by prostaglandins $F_2\alpha$ and E_2 .

VIP is a 28-amino-acid polypeptide originally isolated from porcine duodenum⁶. VIP has subsequently been shown to be present outside the gastrointestinal tract, including in the CNS (central nervous system) where it is concentrated in the cerebral cortex⁷. In this CNS region VIP is contained in a homogeneous population of bipolar neurons that are intracortical and radially

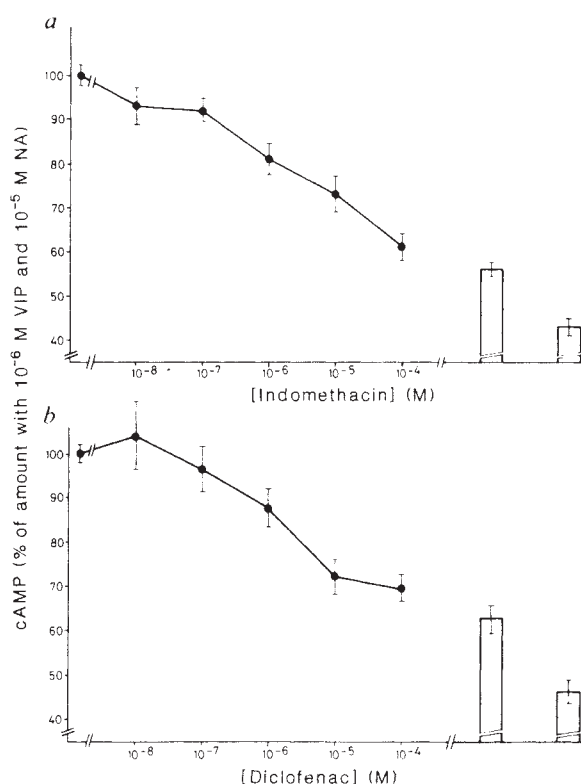


Fig. 1 Concentration-response curve of the antagonism by indomethacin (a) or diclofenac (b) of the synergism between VIP (10^{-6} M) and NA (10^{-5} M) in increasing cAMP levels in mouse cerebral cortical slices. Values are the mean $\pm$ s.e.m. (bars) of 9 to 29 determinations in five (a) and four (b) different experiments. Basal cAMP levels ($\pm$ s.e.m.) were 15.2 ± 2.6 (a) and 3.8 ± 0.5 (b) pmol per mg protein. Absolute levels observed with 10^{-6} M VIP and 10^{-5} M NA were 290.0 ± 13.7 (a) and 204.3 ± 11.7 (b) pmol per mg protein. Throughout, ($\pm$)propranolol was present at 10^{-4} M to antagonize the increases in cAMP elicited by NA through β -adrenergic receptors^{1,2}. The antagonism by indomethacin was statistically significant at 10^{-6} M ($P < 0.001$) and at 10^{-7} M ($P < 0.05$). The antagonism by diclofenac was statistically significant at 10^{-5} M, 10^{-4} M ($P < 0.001$) and at 10^{-6} M ($P < 0.01$). Indomethacin and diclofenac did not affect basal cAMP levels nor the levels observed in presence of 10^{-6} M VIP. Bars (at right) show left, amounts of cAMP in the presence of 10^{-6} M VIP and 10^{-5} M NA and 10^{-5} M of the selective α_1 -antagonist prazosin. This value is not significantly different from cAMP levels observed in the presence of 10^{-6} M VIP and 10^{-5} M NA plus 10^{-4} M indomethacin (in a) or 10^{-4} M diclofenac (in b). Right-hand bar, cAMP levels observed in the presence of 10^{-6} M VIP.

Methods. Cerebral cortical slices were prepared as previously described^{1,2}. The cerebral cortex was dissected on ice and immediately placed in a modified Krebs-Ringer bicarbonate buffer, pH 7.4 (KRG), containing (in mM): NaCl 120, KCl 5, CaCl_2 2.6, MgSO_4 0.67, KH_2PO_4 1.2, glucose 3, NaHCO_3 27.5, previously gassed with O_2/CO_2 (95:5). The dissected cortices (two per experiment) were then placed on a McIlwain tissue chopper, and $250 \times 250 \mu\text{m}$ slices were prepared. The slices were then resuspended in ice-cold KRG (6 ml per cortex) and incubated for 15 min at 37°C . The medium was then replaced and aliquots of the slice suspension (260 μl) were distributed into individual polypropylene test tubes. After gassing with a stream of O_2/CO_2 (95:5), the tubes were capped and incubated at 37°C . After 30 min, drugs (20 μl) were added for 10 min. Propranolol, prazosin, mepacrine, BBPB, TMB-8, indomethacin or diclofenac were added 5 min before the agonists. The slice suspension was then boiled for 10 min and sonicated for 5 s, transferred into Eppendorf test tubes and then centrifuged at 10,000 g for 2 min. Amounts of cAMP were determined as follows. Aliquots of the supernatants (20–200 μl) were lyophilized and resuspended in 20 μl water. Then cAMP was determined in duplicate by the saturation assay described by Brown *et al.*²⁹. Experimental values were calculated by computer from a standard curve ranging from 0–12 pM cAMP. Proteins were then determined in duplicate in the pellets as described by Lowry *et al.*³⁰. VIP was purchased from Professor V. Mutt, Department of Biochemistry II, Karolinska Institutet, Stockholm. NA, PGA_2 , PGD_2 , prostacyclin, PGE_1 , PGE_2 , 6-keto- PGE_1 , melliin, mepacrine, BBPB, TMB-8 were from Sigma. $\text{PGF}_2\alpha$ was a gift of Upjohn Switzerland. Statistical analysis was by Student's *t*-test.

oriented (that is, in a plane perpendicular to the pial surface)⁸. In view of their morphology, VIP-containing neurons are in a position to exert their actions locally, within cortical columns. VIP elicits various cellular actions in the cerebral cortex including the stimulation of cAMP formation^{1,9} and the hydrolysis of

† To whom reprint requests should be addressed.

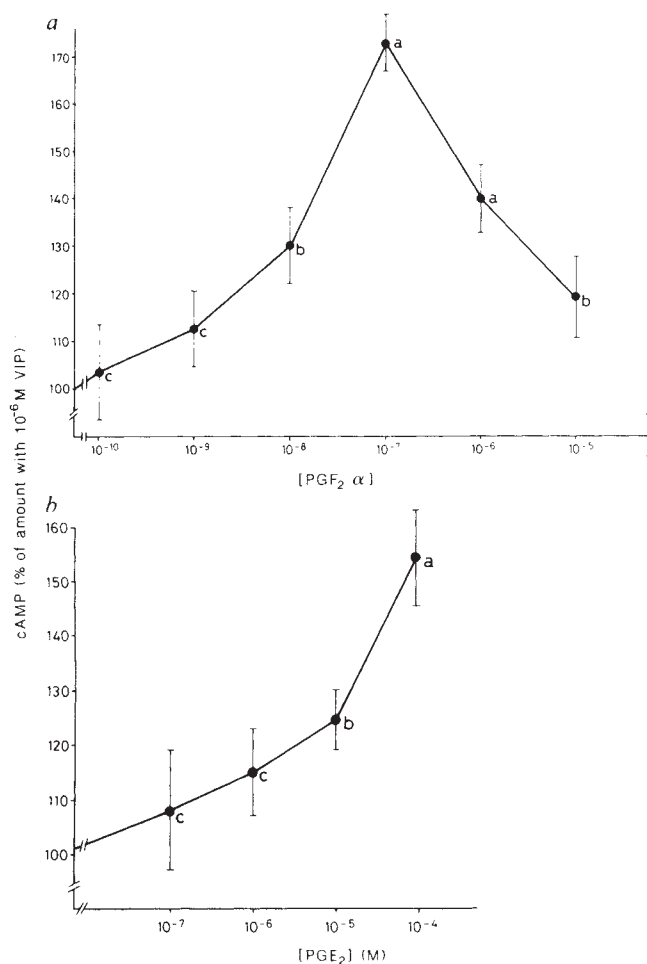


Fig. 2 *a*, Effect of $\text{PGF}_2\alpha$ on cAMP accumulation elicited by 10^{-6} M VIP. For experimental details, see Fig. 1 legend. Data represent the mean $\pm$ s.e.m. of 10–18 determinations in four different experiments. Basal cAMP levels $\pm$ s.e.m. were 10.9 ± 1.1 pmol per mg protein. In the presence of $\text{PGF}_2\alpha$ alone at 10^{-7} M and 10^{-5} M, cAMP levels were 15.7 ± 0.6 and 15.4 ± 1.6 pmoles per mg protein respectively. Absolute levels of cAMP $\pm$ s.e.m. observed with 10^{-6} M VIP were 88.6 ± 4.9 pmol per mg protein. Letters denote significance as follows: a, significantly different from 10^{-6} M VIP ($P < 0.001$); b, significantly different from 10^{-6} M VIP ($P < 0.01$); c, not significantly different from 10^{-6} M VIP. *b*, Effect of PGE_2 on cAMP accumulation elicited by 10^{-6} M VIP. For experimental details, see Fig. 1 legend. Data represent the mean $\pm$ s.e.m. of 4–10 determinations in two different experiments. Basal cAMP levels ($\pm$ s.e.m.) were 10.7 ± 1.2 pmol per mg protein. In the presence of PGE_2 alone at 10^{-7} M and 10^{-5} M, cAMP levels were 9.2 ± 1.2 and 10.2 ± 1.2 pmol per mg protein respectively. Absolute levels of cAMP $\pm$ s.e.m. observed with 10^{-6} M VIP were 142.9 ± 13.0 pmol per mg protein. Significance shown as follows: a, significantly different from 10^{-6} M VIP ($P < 0.001$); b, significantly different from 10^{-6} M VIP ($P < 0.05$); c, not significantly different from 10^{-6} M VIP.

glycogen¹⁰. In view of the latter action, it has been proposed that VIP neurons have a function in the regulation of local cortical energy metabolism¹⁰.

The noradrenergic innervation of the neocortex is provided by fibres originating in the nucleus locus coeruleus in the brainstem¹¹. In the neocortex noradrenergic axons adopt a predominantly tangential trajectory (that is, in a plane parallel to the pial surface), proceeding from the frontal to the occipital poles¹². Thus, in contrast to the locally-acting intracortical VIP-

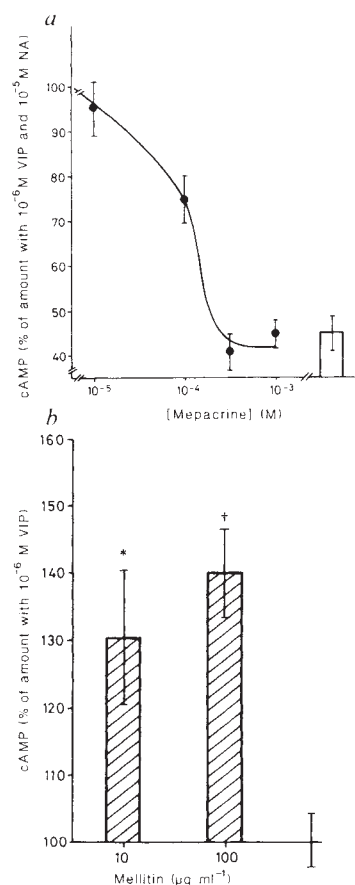


Fig. 3 *a*, Concentration-response curve of the antagonism by mepacrine of the synergism between VIP (10^{-6} M) and NA (10^{-5} M) in increasing cAMP levels in mouse cerebral cortical slices. For experimental details, see Fig. 1 legend. Data represent the mean $\pm$ s.e.m. of 6–22 determinations in four different experiments. Basal cAMP levels ($\pm$ s.e.m.) were 10.8 ± 1.9 (pmol per mg protein). Absolute levels observed with 10^{-6} M VIP + 10^{-5} M NA were 245.9 ± 14.4 (pmol per mg protein). ($\pm$)propranolol was present throughout at 10^{-4} M to antagonize the increases in cAMP elicited by NA through β -adrenergic receptors^{1,2}. The antagonism by mepacrine was statistically significant ($P < 0.001$), except at 10^{-5} M. Mepacrine did not affect basal cAMP levels nor the levels observed in the presence of 10^{-6} M VIP. Bar at right of graph, cAMP levels observed in the presence of 10^{-6} M VIP and 10^{-5} M NA and 10^{-5} M of the selective α_1 -antagonist prazosin. This value is not significantly different from amounts of cAMP observed in presence of 10^{-6} M VIP with 10^{-5} M NA and 5×10^{-4} M mepacrine. *b*, Effect of mellitin on cAMP accumulation elicited by 10^{-6} M VIP. For experimental details, see Fig. 1 legend. In these experiments, bacitracin ($30 \mu\text{g ml}^{-1}$) was added in all the samples 5 min before the adjunction of mellitin and VIP, to protect mellitin from protease activity. Data represent the mean $\pm$ s.e.m. of 11–14 determinations in two different experiments. Basal cAMP levels ($\pm$ s.e.m.) were 3.3 ± 1.0 pmol per mg protein. Levels ($\pm$ s.e.m.) in the presence of mellitin ($100 \mu\text{g ml}^{-1}$) were 2.3 ± 0.8 pmol per mg protein and in the presence of 10^{-6} M VIP 113.0 ± 13.0 pmol per mg protein. * Significantly different from 10^{-6} M VIP ($P < 0.05$); † significantly different from 10^{-6} M VIP ($P < 0.001$).

containing system, noradrenergic fibres are endowed with the capacity to exert their actions globally throughout the neocortex, influencing simultaneously functionally distinct regions¹³. Despite the difference in the morphology of the neuronal systems that contain them, VIP and NA share certain cellular actions in the cerebral cortex, because NA also stimulates cAMP formation and glycogen hydrolysis^{14,15}. In mouse cerebral cortex both effects of NA are mediated by β -adrenergic receptors^{14,15}. In view of the morphological characteristics of the VIP and NA-

Table 1 Effect of TMB-8 on the synergistic interaction between VIP and NA in increasing cAMP levels

Drugs added	Concentration (M)	cAMP (pmol per mg protein)
None		14.0 ± 1.6
VIP*		142.2 ± 10.1
VIP + NA*		
		253.1 ± 14.3
VIP + NA* + TMB-8	10 ⁻⁴	167.1 ± 6.8†
VIP + NA* + TMB-8	10 ⁻⁵	191.5 ± 9.3‡
VIP + NA* + TMB-8	10 ⁻⁶	205.3 ± 14.7§

* 'VIP' and 'NA' refer to concentrations of 10⁻⁶ M and 10⁻⁵ M respectively. Methods are described in Fig. 1 legend. Results are the mean ± s.e.m. of 4–6 separate determinations. Significance shown as follows: † Significantly different from 10⁻⁶ M VIP + 10⁻⁵ M NA ($P < 0.001$); ‡ Significantly different from 10⁻⁶ M VIP + 10⁻⁵ M NA ($P < 0.005$); § Significantly different from 10⁻⁶ M VIP + 10⁻⁵ M NA ($P < 0.05$). TMB-8 at 10⁻⁴ and 10⁻⁵ M did not influence the increases in cAMP levels elicited by 10⁻⁶ M VIP: cAMP levels (pmol per mg protein) were for 10⁻⁵ M TMB-8 + 10⁻⁶ M VIP, 134.1 ± 8.5 ($n = 6$); and 132.7 ± 4.3 ($n = 6$) for 10⁻⁴ M TMB-8 + 10⁻⁶ M VIP.

containing cortical neuronal systems, the synergistic interaction between VIP and NA, mediated by α_1 -adrenergic receptors, should generate marked increases in cAMP levels in discrete cortical volumes ('hot spots')¹. The spatial coordinates of such hot spots would be delineated by the intersection of the tangentially-organized noradrenergic fibres and a group of activated, radially-oriented VIP-containing neurons^{1,2}.

ICR male albino mice (20–25 g) were used throughout the study and cerebral cortical slices were prepared as previously described^{1,2} (see Fig. 1 legend for details). In a first series of experiments we have examined the effect of non-steroidal anti-inflammatory drugs (NSAID) such as indomethacin and diclofenac on the α_1 -adrenergic-mediated synergism between NA and VIP. These drugs are inhibitors of cyclooxygenase, the enzyme responsible for the conversion of arachidonic acid (AA) to prostaglandin (PG) G₂ (ref. 16). As shown in Fig. 1a and b, indomethacin and diclofenac inhibit in a concentration-dependent manner the synergistic interaction between VIP and NA in increasing cAMP levels, with IC₅₀ of $\sim 3 \times 10^{-6}$ M. The α_1 -adrenergic component of the synergism is defined as the increase in cAMP levels elicited by 10⁻⁶ M VIP and 10⁻⁵ M NA that can be inhibited by a maximally effective concentration of the selective α_1 -adrenergic antagonist prazosin, namely 10⁻⁵ M (ref. 2). The selective β -adrenergic antagonist ($\pm$)-propranolol was also present in all conditions at 10⁻⁴ M. This concentration of ($\pm$)-propranolol completely blocks the small increases in cAMP elicited by NA alone without affecting the increases in the cyclic nucleotide stimulated by VIP (ref. 2), nor the synergistic interaction between VIP and NA^{1,2}. As shown in Fig. 1a and b, indomethacin and diclofenac were as effective as prazosin in antagonizing the synergistic interaction between VIP and NA. This observation indicates that inhibition of cyclooxygenase activity and of the formation of AA metabolites interferes with the expression of an α_1 -adrenergic-mediated process. A direct antagonism between indomethacin and NA at the α_1 -adrenergic receptor can be ruled out as we have observed that indomethacin, at the concentrations used for cyclooxygenase inhibition, does not compete with the selective α_1 -adrenergic antagonist ³H-prazosin in radioligand binding experiments (not shown). It should be noted that a small, but statistically significant difference exists between cAMP levels observed with prazosin or cyclooxygenase inhibitors and those found with VIP alone (Fig. 1, far right). This indicates that a small component of the synergistic interaction between VIP and NA is not mediated by α_1 -adrenergic receptors.

In an attempt to determine the nature of the AA metabolites

involved in the synergistic interaction between VIP and NA, we have examined the effect of various prostanoids on the increases in cAMP elicited by VIP alone. Prostaglandin F₂ α (PGF₂ α), potentiated in a concentration-dependent manner the VIP-stimulated cAMP increase, with maximum effect at 10⁻⁷ M (Fig. 2a). Higher concentrations were less effective. PGE₂ was also active, although less potent than PGF₂ α (Fig. 2b). Several other prostanoids, namely prostacyclin, 6-keto-PGE₁, PGE₁, PGA₁ and PGD₂, were tested at concentrations ranging from 10⁻⁹ to 10⁻⁴ M and were inactive, at all concentrations, in stimulating the increases in cAMP elicited by VIP (not shown).

Note that the synergism between VIP and NA could be restored by PGF₂ α after blockade by indomethacin. Results were (in pmol cAMP per mg protein mean ± s.e.m. for four to six observations): VIP at 10⁻⁶ M, 134.9 ± 9.8; VIP at 10⁻⁶ M with NA at 5 × 10⁻⁶ M, 394.8 ± 30.5; VIP at 10⁻⁶ M with NA at 5 × 10⁻⁶ M and indomethacin at 10⁻⁴ M, 222.9 ± 21.9; VIP at 10⁻⁶ M with NA at 5 × 10⁻⁶ M and indomethacin at 10⁻⁴ M and PGF₂ α at 10⁻⁷ M, 373.6 ± 17.3. (The last value is not significantly different from the second).

AA can be formed from membrane phospholipids through the activation of the enzyme phospholipase A₂ (PLA₂)¹⁷. Mepacrine, an inhibitor of PLA₂ (ref. 17), antagonizes in a concentration-dependent manner the α_1 -mediated synergistic interaction between VIP and NA (Fig. 3a). Similar results were obtained with *p*-bromo phenacyl bromide (PBPB) at 10⁻⁵ M, another inhibitor of PLA₂ activity¹⁸ (not shown). As a corollary to these experiments we have observed that mellitin, an activator of PLA₂ (ref. 19) potentiates the increases in cAMP elicited by VIP (Fig. 3b).

The activity of PLA₂ is strongly dependent on the presence of Ca²⁺ (ref. 17). In a series of experiments, we have therefore examined the involvement of intracellular Ca²⁺ in the synergistic interaction between VIP and NA. For these experiments we used 8-(*N,N*-diethylamino)octyl-3,4,5-trimethoxybenzoate-hydrochloride (TMB-8), an inhibitor of Ca²⁺ release from intracellular stores²⁰. The synergistic interaction between VIP and NA was antagonized in a concentration-dependent manner by TMB-8 (Table 1).

These observations suggest that occupation of α_1 -adrenergic receptors by NA could activate PLA₂ and trigger the formation of AA from membrane phospholipids. The AA thus formed would lead, under the action of cyclooxygenase, to the formation of PGF₂ α and/or PGE₂. Similar α_1 -mediated AA formation has been recently demonstrated in a rat thyroid cell line²¹ and prostaglandins of the E series are required for the expression of α -adrenergic-mediated activation of cAMP formation in rat brain slices²². Furthermore, α -adrenergic potentiation of isoprenaline-induced increases in cAMP appear to be associated with PLA₂ activation²³. Note however that α_1 adrenergic receptor activation has been shown to be coupled to another membrane-bound phospholipase, phospholipase C (PLC)²⁴. This coupling seems to involve GTP-binding proteins²⁵. PLC activation results in the formation of diacylglycerol (DAG), an activator of protein kinase C, and of inositol-trisphosphate (InsP₃) which mobilizes Ca²⁺ from intracellular stores, thus increasing Ca²⁺ concentrations²⁴. AA formation may result from stimulation of PLC activity²⁶ in addition to PLA₂ activation. Thus AA can be formed from DAG by the enzyme DAG lipase²⁷. Furthermore, the increases in intracellular Ca²⁺ elicited by the PLC-mediated InsP₃ formation may result in the activation of PLA₂, which is dependent on Ca²⁺ for its activity¹⁷. PLA₂ activation and AA formation could therefore also result from α_1 -mediated PLC activation by a number of sequentially-linked mechanisms.

Our observations indicate that an α_1 -mediated action of NA in the cerebral cortex, namely potentiation of the increases in cAMP elicited by VIP, involves the formation of AA metabolites, possibly PGF₂ α and PGE₂. The formation of AA and of its metabolites seems to result from the activation of PLA₂, as

indicated by the antagonism of the synergism by mepacrine and PBPB, by the action of the PLA₂ activator mellitin, and by the sensitivity to the inhibition of intracellular Ca²⁺ release by TMB-8. The activation of PLA₂ may in turn result either from a direct coupling between α_1 -adrenergic receptors and PLA₂, possibly through a GTP-binding protein as suggested for the FRTL5 thyroid cell line²⁸, or from indirect mechanisms mediated by PLC, following activation of α_1 -adrenergic receptors.

This work was supported by Fonds National Suisse de la Recherche Scientifique (FNRS) grants to P.J.M. and to M.S. P.J.M. is the recipient of a Career Development Award (START) from FNRS. N.C.S. is a recipient of Fellowships from Fondation Suisse de Recherche sur l'Alcool and Fondation Pierre Mercier. We thank Ms Sylvianne Bonnet for expert secretarial help, Mr Fred Pillonel for graphical work, and Drs P. L. Herrling and S. Urwiler, Wander AG Forschungsinstitut, Bern, for help with the ³H-prazosin binding assay.

Received 24 March; accepted 2 June 1987.

1. Magistretti, P. J. & Schorderet, M. *Nature* **308**, 280-282 (1984).
2. Magistretti, P. J. & Schorderet, M. *J. Neurosci.* **5**, 362-368 (1985).
3. Hollingsworth, E. B. & Daly, J. W. *Biochim. biophys. Acta* **847**, 207-216 (1985).
4. Redgate, E. S., Deupree, J. D. & Axelrod, J. *Brain Res.* **365**, 61-69 (1986).

5. Ferron, A., Siggins, G. R. & Bloom, F. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8810-8812 (1985).
6. Said, S. I. & Mutt, V. *Science* **169**, 1217-1218 (1970).
7. Lorén, I. et al. *Neuroscience* **4**, 1953-1976 (1979).
8. Morrison, J. H., Magistretti, P. J., Benoît, R. & Bloom, F. E. *Brain Res.* **292**, 269-282 (1984).
9. Quick, M., Iversen, L. L. & Bloom, S. R. *Biochem. Pharmac.* **27**, 2209-2213 (1978).
10. Magistretti, P. J., Morrison, J. H., Shoemaker, W. J., Sapin, V. & Bloom, F. E. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6535-6539 (1981).
11. Dahlstrom, A. & Fuxe, K. *Acta Physiol. scand.* **62** (suppl. 232), 1-55 (1964).
12. Morrison, J. H., Molliver, M. E., Grzanna, R. & Coyle, J. T. *Neuroscience* **6**, 139-158 (1981).
13. Morrison, J. H. & Magistretti, P. J. *Trends Neurosci.* **6**, 146-151 (1983).
14. Daly, J. in *Handbook of Psychopharmacology* (eds Iversen, L. L., Iversen, S. D. & Snyder, S. H.) 47-129 (Plenum, New York, 1975).
15. Quach, T. T., Rose, C. & Schwartz, J. C. *J. Neurochem.* **30**, 1335-1341 (1978).
16. Vane, J. R. *Nature new Biol.* **231**, 232-235 (1971).
17. Irvine, R. F. *Biochem. J.* **204**, 3-16 (1982).
18. Roberts, M. F., Deems, R. A., Mincey, T. C. & Dennis, E. A. *J. biol. Chem.* **252**, 2405-2411 (1977).
19. Shier, W. T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 195-199 (1979).
20. Chiou, C. Y. & Malagodi, M. H. *Br. J. Pharmac.* **53**, 279-285 (1975).
21. Burch, R. M. et al. *J. biol. Chem.* **261**, 11236-11241 (1986).
22. Partington, C. R., Edwards, M. W. & Daly, J. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3024-3028 (1980).
23. Duman, R. S., Karbon, E. W., Harrington, C. & Enna, S. J. *J. Neurochem.* **47**, 800-810 (1986).
24. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315-321 (1984).
25. Cockcroft, S. & Gomperts, B. D. *Nature* **314**, 534-536 (1985).
26. Lapetina, E. G. *Trends Pharmac. Sci.* **3**, 115-118 (1982).
27. Bell, R. L., Kennerly, D. A., Stanford, N. & Majerus, P. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3238-3241 (1979).
28. Burch, R. M., Luini, A. & Axelrod, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7201-7205 (1986).
29. Brown, B. L., Albano, J. D. M., Ekins, R. P., Sghezzi, A. M. & Tampion, W. *Biochem. J.* **121**, 561-562 (1971).
30. Lowry, O. H., Rosenbrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

Micromolar concentrations of Zn²⁺ antagonize NMDA and GABA responses of hippocampal neurons

Gary L. Westbrook & Mark L. Mayer

Laboratory of Developmental Neurobiology,
National Institute of Child Health and Human Development,
National Institutes of Health, Bethesda, Maryland 20892, USA

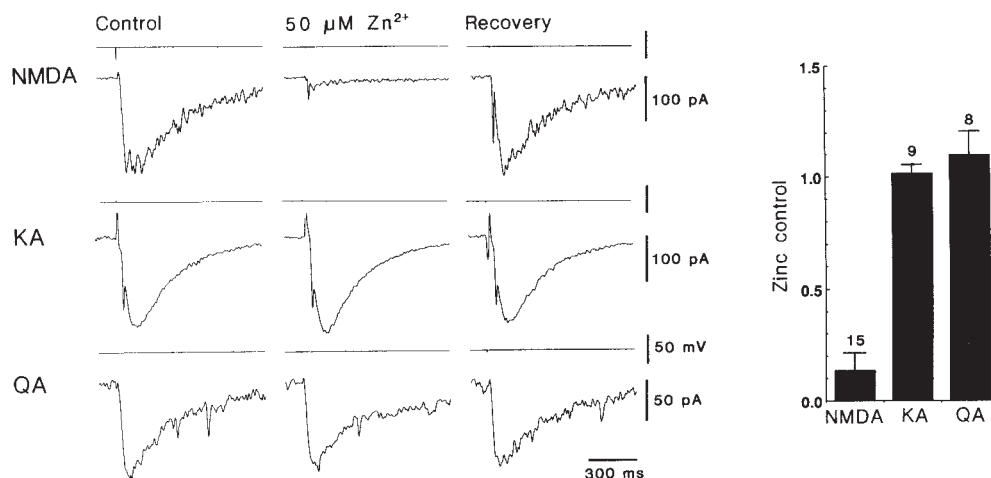
NMDA (*N*-methyl-D-aspartate) receptors serve as modulators of synaptic transmission in the mammalian central nervous system (CNS) with both short-term and long-lasting effects¹. Divalent cations are pivotal in determining this behaviour in that Mg²⁺ blocks the ion channel in a voltage-dependent manner^{2,3}, and Ca²⁺ permeates NMDA channels⁴. Zn²⁺ could also modulate neuronal excitability because it is present at high concentrations in brain, especially the synaptic vesicles of mossy fibres in the hippocampus^{5,6} and is released with neuronal activity^{7,8}. Both proconvulsant⁹ and depressant¹⁰ actions of Zn²⁺ have been reported. We have found that zinc is a potent non-competitive antagonist of NMDA responses on cultured hippocampal neurons. Unlike Mg²⁺, the effect of Zn²⁺ is not voltage-sensitive between -40 and +60 mV, suggesting that Zn²⁺ and Mg²⁺ act at distinct sites. In addition, we have found that Zn²⁺ antagonizes responses to the inhibitory transmitter GABA (γ -aminobutyric acid). Our results provide evidence for an additional metal-binding site on the NMDA receptor channel, and suggest that Zn²⁺ may regulate both excitatory and inhibitory synaptic transmission in the hippocampus.

Receptors for excitatory amino acids can be distinguished on the basis of responses to the selective agonists NMDA, kainic acid and quisqualic acid^{1,11}. The effect of zinc on responses to each of these agonists was examined using hippocampal neurons grown in dissociated cell culture. Zn²⁺ (50 μ M) produces a nearly complete block of the NMDA-activated conductance on hippocampal neurons voltage clamped near their resting membrane potential; in contrast, responses to kainic acid and quisqualic acid are slightly potentiated (Fig. 1). The peak amplitude of agonist responses compared to control was NMDA, 0.14 ± 0.08 (mean $\pm$ s.d., $n = 15$); QA, 1.1 ± 0.1 , $n = 8$; KA, 1.02 ± 0.09 , $n = 9$ (Fig. 1). These effects were significantly different from the

control response for NMDA ($P < 0.001$, one-tailed t test) and quisqualic acid ($P < 0.05$). Higher concentrations of Zn²⁺ produced a greater potentiation of responses to kainate and quisqualate and, at 250 μ M Zn²⁺, complete block of responses to NMDA. The other Group IIB metals, cadmium and mercury, also blocked responses to NMDA and appeared to be similar in action to Zn²⁺, although cadmium was approximately 50-fold less potent than zinc.

Previous work with zinc suggests two possible mechanisms for its action as an NMDA antagonist: (1) reaction of sulphhydryl bonds in the receptor-channel protein with Zn²⁺, a property shared by mercury and cadmium, or (2) formation of inactive complexes between Zn²⁺ and either glycine, a potent modulator of NMDA receptor-channel activity¹², or NMDA itself. However neither of these mechanisms appears to explain the antagonism of NMDA responses by Zn²⁺. First, the effect of Zn²⁺ was rapidly and completely reversible (Fig. 1) which is inconsistent with chemical modification of sulphhydryl bonds. Likewise, Zn²⁺ blocked NMDA responses at concentrations much lower than required for reaction of zinc with sulphhydryl bonds in other proteins¹³⁻¹⁵. This contrasts with the irreversible reduction of the response of acutely dissociated hippocampal neurons to kainate by micromolar Hg²⁺ or 5-10 mM zinc¹³. We have confirmed this irreversible action of Hg²⁺ on the response to kainate, whereas in the same experiments responses to NMDA showed rapid recovery. Second, glycine, acting at a strychnine-insensitive site with a nanomolar affinity constant, causes a marked potentiation of NMDA responses¹², but in our experiments potentiation of NMDA responses by a supramaximal concentration of glycine (20 μ M) did not protect against Zn²⁺ antagonism (two neurons, seven applications of 10 μ M Zn²⁺, 52% reduction). This suggests that the effects of Zn²⁺ cannot be explained by the formation of inactive complexes with glycine, and that glycine-modulated NMDA receptors are also Zn²⁺-sensitive. The high potency of Zn²⁺ is also unlikely to reflect chelation of NMDA itself. Calculations using stability coefficients published for aspartic acid (with corrections for pH)¹⁶ indicate that with 10 μ M aspartic acid and 50 μ M Zn²⁺ complex formation will only reduce the concentration of free aspartic acid to around 9 μ M at pH 7.4. The affinity of zinc for glycine is even less than for aspartic acid, further suggesting that complex formation between Zn²⁺ and glycine is an unlikely explanation for our results. Thus a direct action on the receptor channel complex seems probable.

Fig. 1 NMDA responses are selectively and potently antagonized by Zn^{2+} . Left, inward currents evoked by NMDA (20 ms, 130 nA) were rapidly and reversibly blocked by Zn^{2+} whereas responses to kainic acid (KA: 35 ms, 200 nA) and quisqualic acid (QA: 15 ms, 100 nA) were slightly potentiated. Each set of records shows the membrane potential voltage clamped at -60 mV, and the current response to an iontophoretic pulse of excitatory amino acid. Responses to NMDA and kainate were recorded from the same neuron; the increase in current noise evoked by NMDA was typical, and reflects the open-shut transitions underlying ion channel gating. Right, histogram showing that the antagonism by Zn^{2+} (expressed as a fraction



of the control response, mean $\pm$ s.d.) was highly selective for the NMDA receptor subtype. Numbers above each bar, number of neurons tested.

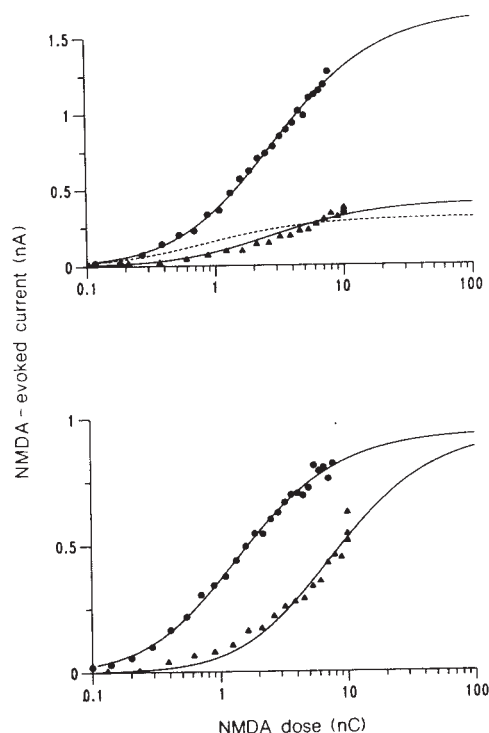
Methods. Low-density cultures of hippocampal neurons were prepared from 17-day mouse embryos by plating neurons onto confluent layers of astrocytes prepared from the hippocampal region. Methods of dissociation and maintenance of cultures were similar to those described before²⁰ except that the neuronal cultures were treated on day 1 with antimitotics to suppress overgrowth of background cells. Cultures were maintained in a medium containing 5% horse serum (Hyclone), 95% Minimal Essential Medium (MEM, Gibco) and a nutrient supplement. The concentration of Zn^{2+} in this medium (contributed by the serum) was estimated at 0.5μ M. Before recording, the plates were washed with a salt solution containing (mM): NaCl 145, KCl 2.5, $CaCl_2$ 1, HEPES 10, 1μ M TTX and 100μ M picrotoxin with no added Mg^{2+} . Neurons were voltage clamped using CsCl-filled patch electrodes in the whole-cell recording mode, and an Axon Instruments Axoclamp-2 amplifier switching at 10–14 kHz. The patch pipette solution contained (in mM) 140 CsCl, 10 HEPES, 2 $MgCl_2$, 10 glucose and 1.1 EGTA, titrated to pH 7.2 with CsOH. NMDA (0.2 M, pH 8.0), kainic acid (0.5 M, pH 8.0) and quisqualic acid (0.005 M in 0.1 M NaCl at pH 8.0) were applied by iontophoresis from microelectrodes (80–100 M Ω) positioned within 5μ m of the soma. Zn^{2+} (50μ M) was diluted in the recording solution and applied from a patch pipette using a pressure pulse terminating 180 ms before the agonist application. Experiments were performed at room temperature (25 – 27° C).

Fig. 2 Above, Zn^{2+} is a non-competitive NMDA antagonist.

Dose-response curves were constructed using iontophoretic pulses of 20 ms duration, intensity 4.8–500 nA, at a holding potential of -60 mV. Control dose-response curves were fitted using a non-linear least-squares fitting program to an equation of the form, $I = I_{\max(\text{control})} \times ([A]/([A] + K_D))^n$ where $[A]$ is the dose of NMDA measured in nanocoulombs and the number of sites (n), the maximum response ($I_{\max(\text{control})}$) and agonist affinity constant (K_D) were allowed to vary to obtain the best fit. In experiments on 8 neurons (26 dose-response curves) the value obtained for n varied from 0.9 to 2.4 (1.55 ± 0.48 , mean $\pm$ s.d.). Dose-response curves recorded during application of Zn^{2+} and AP5 were fitted using the estimates for n , $I_{\max(\text{control})}$ and K_D from the control dose-response curve fit for each neuron, and allowing the antagonist affinity constant (K_i) to vary to obtain the best fit. Below, the antagonist DL-AP5 (30μ M) causes a rightward shift of the iontophoretic dose-response curve for the NMDA-activated conductance (control values were $n = 1.7$; $K_D = 0.76$ nC; $I_{\max(\text{control})} = 0.9$ nA) which could be fitted with a competitive model where binding of agonist and antagonist are assumed to be mutually exclusive:

$$I = I_{\max(\text{control})} \times \left\{ \frac{[A]}{[A] + (K_D \times (1 + [i]/K_i))} \right\}^n$$

However, antagonism by Zn^{2+} (40μ M) was clearly non-competitive (control values of upper graph, $n = 1.4$; $K_D = 1.94$ nC; $I_{\max(\text{control})} = 1.7$ nA) where $I = I_{\max(\text{control})} \times ([A]/([A] + K_D))^n \times (1 - ([i]/([i] + K_i)))$. Broken line, the best fit obtained using an uncompetitive model of antagonism where $I = \frac{I_{\max(\text{control})} \times ([A]/([A] \times (1 + ([i]/K_i)) + K_D))^n}{1 + ([i]/K_i)}$. The non-competitive model assumes agonist and antagonist bind independently at different sites, and thus the antagonist reduces $I_{\max}$ without changing K_D and the degree of antagonism does not vary with changes in $[A]$. In the uncompetitive model the antagonist can only bind to the agonist-receptor complex and thus the antagonism increases with increasing $[A]$. For every neuron examined competitive, non-competitive and uncompetitive models were fitted to the data for both Zn^{2+} and AP5; the best fit for Zn^{2+} was always obtained using a non-competitive model of antagonism, whereas for AP5 the best fit was obtained using a competitive model of antagonism. Zn^{2+} and AP5 were applied by continuous local perfusion from a small-tipped (2μ m) puffer pipette positioned 50μ m from the soma, thus the concentration of antagonists at the neuronal membrane was considerably less than that in the pipettes. Data from two neurons.



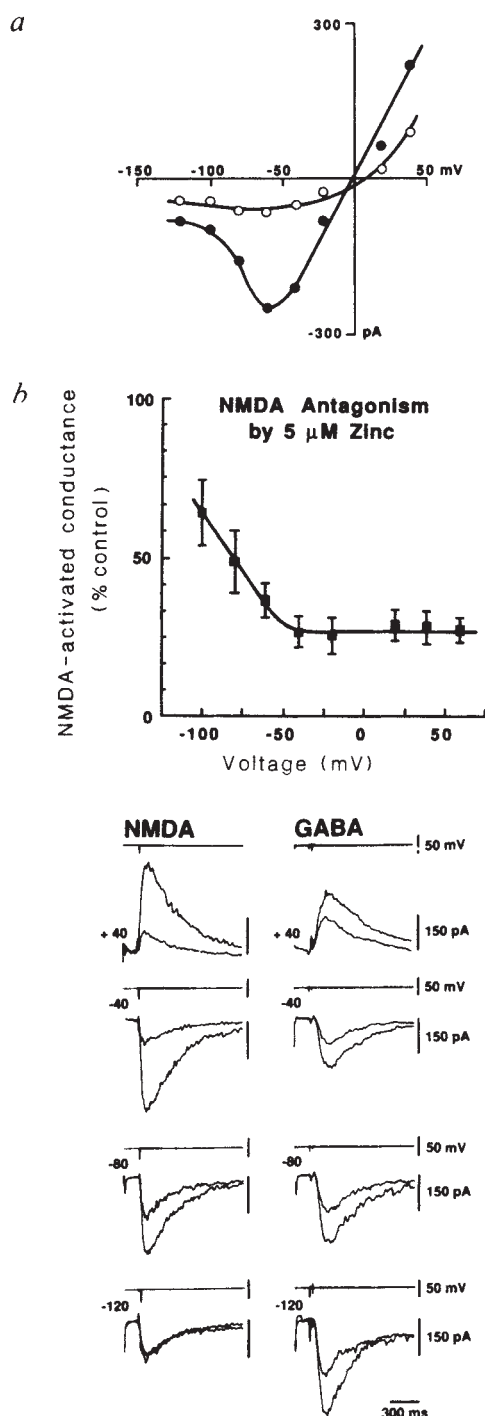


Fig. 3 Zn^{2+} blocks the NMDA receptor-channel at a site distinct from Mg^{2+} . In **a**, current-voltage (I - V) relationships for NMDA-activated currents are shown before ($\bullet$) and during ($\circ$) application of $5 \mu\text{M}$ Zn^{2+} . The control I - V plot shows a negative slope conductance due to residual Mg^{2+} (estimated at 3 - $5 \mu\text{M}$). In the presence of Zn^{2+} the NMDA-activated conductance was reduced over a wide range of membrane potentials. Unlike Mg^{2+} , Zn^{2+} blocked outward current flow through NMDA receptor channels at membrane potentials positive to the reversal potential of 0 mV. Methods are given in detail in Fig. 1 legend. **b**, Plot of the data illustrated in **a** for 16 neurons, presented as the block of the NMDA-activated conductance produced by $5 \mu\text{M}$ Zn^{2+} plotted against membrane potential (expressed as mean $\pm$ s.d. for the number of observations). The number of observations ranged from $n = 42$ at -60 mV to $n = 13$ at -100 and $+40$ mV. The apparent decrease in the effect of zinc at membrane potentials negative to -60 mV contrasts with the action of Mg^{2+} . **c**, Zn^{2+} also antagonizes chloride-dependent responses to GABA. Shown are the effects of Zn^{2+} ($5 \mu\text{M}$) on responses to iontophoretic applications of NMDA and GABA recorded from one neuron; current records in the presence and absence of Zn^{2+} are superimposed at several holding potentials. The degree of antagonism of GABA by Zn^{2+} was similar over a wide range of membrane potentials.

Zn^{2+} could compete at the agonist binding site, allosterically modify receptor-channel coupling (like glycine¹²) or block the channel (like Mg^{2+} , refs 2 and 3). Millimolar concentrations of Zn^{2+} interfere directly with the binding of L-glutamate to its receptors¹⁷. To distinguish between these possibilities, we compared the behaviour of Zn^{2+} with that of the competitive NMDA receptor antagonist 2-amino-5-phosphonopentanoic acid (AP5)^{18,19} on currents evoked by iontophoretically applied NMDA (Fig. 2). As expected, $30 \mu\text{M}$ ($\pm$)AP5 produced a rightward shift of the dose-response curve (Fig. 2, $n = 4$). However, antagonism by Zn^{2+} was clearly not competitive ($n = 5$). Using a nonlinear least-squares fitting routine, Zn^{2+} antagonism could be best described by a non-competitive model (see Fig. 2 legend).

Mg^{2+} binds to a site within the NMDA receptor channel, producing a voltage-dependent block which increases at hyperpolarized membrane potentials, and which shows relief with depolarization^{2,3}. This behaviour is consistent with an action as a positively charged channel blocker. Similar effects are also seen with several other divalent cations including Ni^{2+} , Co^{2+} , and Mn^{2+} (refs 20 and 21). However, the effectiveness of Zn^{2+} was unaltered as the membrane potential was depolarized over the range -40 mV to $+60$ mV (Fig. 3a and b). In the presence of micromolar Mg^{2+} , the potency of zinc seemed to decrease as the membrane potential was hyperpolarized beyond -40 mV, so that at -120 mV zinc was in some experiments ineffective as an NMDA antagonist (Fig. 3b and c). Thus it is extremely unlikely that Zn^{2+} and Mg^{2+} act a common site. These results also suggest that the Zn^{2+} binding site is near the external face of the membrane rather than deep within the channel as has been suggested for the Mg^{2+} binding site^{2,20}.

Antagonism of NMDA responses seems an unlikely explanation for the proconvulsant action of Zn^{2+} . Previous studies reported that Zn^{2+} reduced responses to GABA on lobster muscle²², but on vertebrate neurons Zn^{2+} either had no effect²² or potentiated²³ responses to GABA. In our experiments, chloride-dependent responses to GABA were reduced by $5 \mu\text{M}$ Zn^{2+} (peak response at -60 mV reduced to 0.55 ± 0.08 of control, $n = 3$) over a wide range of membrane potentials (Fig. 3c). This action of Zn^{2+} on responses to GABA was of lower potency than recorded for block of NMDA receptors, because on the same cells responses to NMDA were reduced to 0.36 ± 0.11 of control. In cone photoreceptors from turtle retina, several exogenous divalent cations including Co^{2+} , Ni^{2+} and Cd^{2+} (at $10 \mu\text{M}$) also produce a voltage-independent block of GABA responses²⁴.

Our results suggest the presence of a third regulatory site on the NMDA receptor channel complex in addition to those for Mg^{2+} and glycine. The molecular mechanisms at these sites will be of interest for future studies probing structure-function relationships for this ion channel. Although Zn^{2+} , Cd^{2+} and Hg^{2+} all react with sulphhydryl bonds, this seems unlikely to underlie the action of Zn^{2+} on NMDA receptor channels. Zn^{2+} is known to coordinate with other side-chain residues in proteins, including the imidazole ring in histidine, but further work will be required to resolve its site of action at a molecular level. The functional significance of Zn^{2+} antagonism of NMDA and GABA responses in the hippocampus deserves further consideration given that synaptic Zn^{2+} release has been estimated to achieve a far higher concentration⁷ than that necessary to antagonize NMDA and GABA responses. Of importance will be an understanding of the free compared to the bound concentrations of Zn^{2+} in the synaptic cleft, because Zn^{2+} is avidly chelated by amino acids and proteins. The effects of Zn^{2+} described here contrast with the toxic effects of millimolar Zn^{2+} on cultured cortical neurons²⁵; but such high tonic extracellular concentrations are unlikely to occur during synaptic transmission. Our observations suggest that any postsynaptic action of Zn^{2+} on neuronal excitability will be influenced by the balance of excitatory and inhibitory inputs. Regional differences in NMDA-receptor density²⁶ also might influence the effects of

Zn^{2+} on excitability; long-term potentiation of commissural but not mossy fibre inputs to CA3 is blocked by NMDA receptor antagonists²⁷. This may reflect either a lower density of NMDA receptor sites in stratum lucidum where mossy fibres synapse, of the release of Zn^{2+} and block of NMDA receptors during synaptic transmission. Zn^{2+} released from mossy fibre synapses might also under some circumstances influence transmission at adjacent synapses: for example activation of NMDA receptors by commissural afferents might be reduced during stimulation of mossy fibres.

Note added in proof: After submission of this manuscript and publication of a preliminary communication²⁸, selective antagonism by zinc of NMDA responses in cortical neurones was also described by Peters *et al.*²⁹.

Received 20 April; accepted 18 June 1987.

1. Mayer, M. L. & Westbrook, G. L. *Prog. Neurobiol.* **28**, 197-276 (1987).
2. Nowak, L., Bregestovski, P., Ascher, P., Herbet, A. & Prochiantz, A. *Nature* **307**, 462-465 (1984).
3. Mayer, M. L., Westbrook, G. L. & Guthrie, P. B. *Nature* **309**, 261-263 (1984).
4. MacDermott, A. B., Mayer, M. L., Westbrook, G. L., Smith, S. J. & Barker, J. L. *Nature* **321**, 519-522 (1986).

5. Frederickson, C. J., Klitenick, M. A., Manton, W. I. & Kirkpatrick, J. B. *Brain Res.* **273**, 335-339 (1983).
6. Perez-Clausell, J. & Danscher, G. *Brain Res.* **337**, 91-98 (1985).
7. Assaf, S. Y. & Chung, S. H. *Nature* **308**, 734-736 (1984).
8. Howell, G. A., Welch, M. G. & Frederickson, C. J. *Nature* **308**, 736-738 (1984).
9. Chung, S. H., Gabrielson, B. & Norris, D. K. *Advances in Experimental Medicine and Biology* Vol. 203: *Excitatory Amino Acids and Epilepsy* (eds Schwarcz, R. & Ben-Ari, Y.) 545-555 (Plenum, New York, 1986).
10. Khulusi, S. S., Brown, M. W. & Wright, M. *Brain Res.* **363**, 152-155 (1986).
11. Watkins, J. C. & Evans, R. H. A. *Rev. Pharmac. Tox.* **21**, 165-204 (1981).
12. Johnson, J. W. & Ascher, P. *Nature* **325**, 529-531 (1987).
13. Kiskin, N. I., Khrishtal, O. A., Tsyndrenko, A. Ya. & Akaike, N. *Neurosci. Lett.* **66**, 305-310 (1986).
14. Li, N. C. & Manning, R. A. *J. Am. chem. Soc.* **77**, 5225-5228 (1955).
15. Simpson, R. B. *J. Am. chem. Soc.* **83**, 4711-4717 (1961).
16. O'Sullivan, W. J. in *Data for Biochemical Research* (eds Dawson, R., Elliott, D., Elliott, W. & Jones, K.) 423-434 (Clarendon, Oxford, 1969).
17. Mena, E. E., Monaghan, D. T., Whittemore, S. R. & Cotman, C. W. *Brain Res.* **329**, 319-322 (1985).
18. Olverman, H. J., Jones, A. W. & Watkins, J. C. *Nature* **307**, 460-462 (1984).
19. Harrison, N. L. & Simmonds, M. A., *Br. J. Pharmac.* **84**, 381-391 (1985).
20. Mayer, M. L. & Westbrook, G. L. *J. Physiol., Lond.* **361**, 65-90 (1985).
21. Mayer, M. L. & Westbrook, G. L. *J. Physiol., Lond.* (in the press).
22. Smart, T. G. & Constanti, A. *Proc. R. Soc. B* **215**, 327-341 (1982).
23. Smart, T. G. & Constanti, A. *Neurosci. Lett.* **40**, 205-211 (1983).
24. Kaneko, A. & Tachibana, M. *J. Physiol., Lond.* **373**, 463-479 (1986).
25. Yokoyama, M., Koh, J. & Choi, D. W. *Neurosci. Lett.* **71**, 351-355 (1986).
26. Monaghan, D. T. & Cotman, C. W. *J. Neurosci.* **5**, 2909-2919 (1985).
27. Harris, E. W. & Cotman, C. W. *Neurosci. Lett.* **70**, 132-137 (1986).
28. Mayer, M. L. & Westbrook, G. L. *Biophys. J.* **51**, 64a (1987).
29. Peters, S., Koh, J. & Choi, D. W. *Science* **236**, 589-593 (1987).

Changes in hypothalamic preproenkephalin A mRNA following stress and opiate withdrawal

Stafford L. Lightman* & W. Scott Young III

Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, Maryland 20892, USA

The median eminence of the pituitary is rich in opioid receptors¹, and exogenous opioids have major effects on the release of adrenocorticotrophic hormone (ACTH)^{2,3}, luteinizing hormone (LH)^{4,5}, prolactin^{4,6}, growth hormone (GH)^{4,6} and thyrotropin^{7,8}. Stress results in similar changes in anterior pituitary hormone secretion. Enkephalin immunoreactivity has been reported in the medial parvocellular neurons of the hypothalamic paraventricular nucleus which project to the median eminence⁹⁻¹², the site where hypothalamic releasing factors are secreted into the portal blood and thence to the anterior pituitary gland. The endocrine response to stressful stimuli might therefore, at least in part, be mediated through the activation of hypothalamic enkephalinergic neurons. We show that two stressful stimuli, opiate withdrawal and intraperitoneal injection of hypertonic saline, both result in very rapid and marked increases in enkephalin mRNA in the parvocellular paraventricular nucleus. The activation of hypothalamic enkephalin neurons may be important in the neuroendocrine response to stress.

There is surprisingly little information about the physiological control and function of endogenous opioid systems in the central nervous system and, in spite of reported effects of exogenous opioids on anterior pituitary function, it is not known whether the hypothalamic enkephalin system can be modified by neuroendocrine events. We have studied the response of hypothalamic enkephalin messenger RNA to two stressful stimuli in 200-g male Sprague-Dawley rats.

One group of rats was treated with subcutaneous morphine, naloxone or saline vehicle. After 12 days the subcutaneous pumps were removed under ether anaesthesia from both control

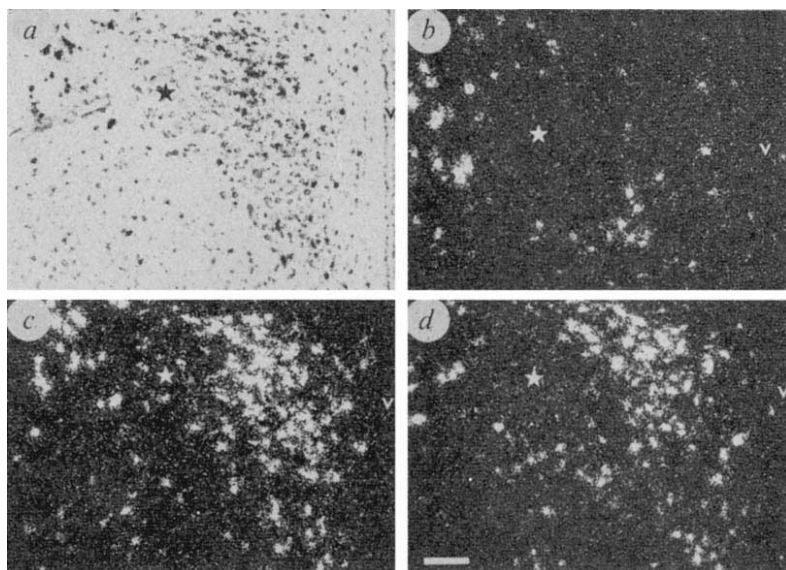
and morphine-treated animals. Two hours later (time 0) and at 3, 7 and 23 hours, naloxone (3 mg per kg body weight) was injected subcutaneously. A separate group of rats was stressed by intraperitoneal injection of 1.5 M NaCl (1.8 ml per 100 g).

Enkephalin immunoreactivity has been reported in the medial parvocellular subdivision of the paraventricular nucleus⁹⁻¹², but in normal animals we could detect only a few scattered cells in this area containing enkephalin mRNA (Fig. 1b). Many cells strongly labelled with the enkephalin probe were detected in other areas, including the striatum and the perifornical area of the hypothalamus. Neither morphine alone nor naloxone had any significant effect on this distribution of enkephalin mRNA. In contrast, naloxone-precipitated opiate withdrawal resulted in a massive and extremely rapid accumulation of enkephalin mRNA in cells of the parvocellular paraventricular nucleus within 2 h, and the level remained high at 24 h (Figs 1 and 2). There was also a similar response to the stress of intraperitoneal (i.p.) hypertonic saline (Figs 1 and 2). The effect was specific to the medial parvocellular area of the paraventricular nucleus: background readings taken adjacent to this area remained very low and no change in enkephalin hybridization occurred in either the striatum or perifornical area (striatum: control, 0.24 ± 0.09 ; N12, 0.27 ± 0.11 ; M1, 0.26 ± 0.08 ; M12, 0.27 ± 0.09 ; W4, 0.25 ± 0.09 ; W8, 0.26 ± 0.12 ; and W24 0.24 ± 0.08 copies of probe bound per μm^3 of tissue; perifornical area: control, 0.11 ± 0.03 ; N12, 0.1 ± 0.02 ; M1, 0.09 ± 0.03 ; M12, 0.11 ± 0.03 ; W4, 0.14 ± 0.05 ; W8, 0.12 ± 0.03 ; and W24, 0.12 ± 0.04 copies of probe bound per μm^3 of tissue. Values are means $\pm$ s.d.; for the key to the treatment codes see Fig. 2 legend).

Our data show a dramatic and extremely rapid accumulation of mRNA in the enkephalin-containing parvocellular neurons of the paraventricular nucleus. There is immunocytochemical evidence for the coexistence of enkephalin and corticotropin releasing factor (CRF) both in parvocellular paraventricular cells⁹ and in granules in the median eminence¹³. We have also found a marked increase in CRF mRNA in response to stress with a similar time course to that for enkephalin mRNA (unpublished data). There is evidence that these parvocellular paraventricular cells project to the median eminence^{13,14} suggesting that their activation may contribute to the complex neuroendocrine consequences of stress (Fig. 3). Enkephalin released from parvocellular paraventricular neurons may have direct actions at the median eminence on terminals containing hypophysiotropic releasing factors or may modulate the activity of

* Present address: Medical Unit, Charing Cross and Westminster Medical School, Westminster Hospital, 17 Page Street, London SW1P 2AP, UK.

Fig. 1 Distribution of autoradiographic grains which represent enkephalin probe hybridized in cells of the paraventricular nucleus (PVN) of the hypothalamus. The silver grains are generated by dipping the slides in NTB3 emulsion and exposure for 35 days. The parvocellular neurons of the PVN are found between the magnocellular core of the nucleus (asterisk) and the third ventricle (v). Labeled perifornical enkephalin cells (to the left of the asterisk) are also seen in all conditions. *a*, Brightfield photomicrograph 24 h after naloxone-precipitated morphine withdrawal. *b-d*, Darkfield photomicrographs: *b*, control animals; *c*, 24 h after naloxone-precipitated withdrawal (darkfield of *a*); *d*, 4 h after i.p. hypertonic saline. Control animals have very few parvocellular cells containing detectable enkephalin transcripts. Opiate withdrawal and hypertonic saline result in a marked increase both in total enkephalin mRNA (see Fig. 2) and the number of cells containing detectable message. Scale bar, 100 μ m.



Methods. Rats treated with opiates received subcutaneous morphine (8 mg per kg) twice daily on day 1 and 16 mg per kg at 09:00 on day 2 and then at 18:00 on the same day, osmotic minipumps (Alzet 2ML1; Alza, Palo Alto, USA) delivering 60 mg per kg daily were implanted subcutaneously under ether anaesthesia. A group of naloxone-treated animals received 5 mg per kg subcutaneous (s.c.) naloxone on day 1 and on day 2 were implanted with osmotic minipumps delivering 20 mg per kg per day. Control groups received vehicle injections (0.9% saline) and were implanted with spent minipumps. Brains were removed rapidly, frozen on dry ice and stored at -40°C . Tissue processing was performed according to Young *et al.*²⁶, on frozen sections (12 μ m) cut 1.8–1.9 mm caudal to the bregma. The enkephalin probe was directed against bases 388–435 (ref. 27) and was made on an Applied Biosystems DNA Synthesizer (Dr M. Brownstein, NIMH). Confirmation of specificity was obtained by Northern analysis using total RNA from rat hypothalamus. The enkephalin probe hybridized to a single band of 1,500 bases under an identical degree of stringency as that used in the current *in situ* studies. The probe was labelled using ^{35}S -dATP ($>1,000\text{ Ci nmol}^{-1}$, NEN) and terminal deoxynucleotidyl transferase to a specific activity of $5.6 \times 10^8\text{ d.p.m. } \mu\text{g}^{-1}$.

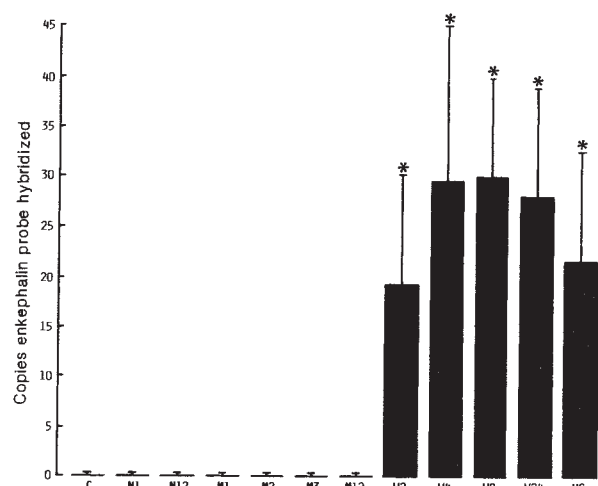


Fig. 2 Quantitative *in situ* hybridization histochemical estimation of hypothalamic parvocellular paraventricular mRNA for enkephalin. Results expressed as copies of probe hybridized per paraventricular nucleus per section ($\times 10^4$). C, control; N1, 4 h after injection of naloxone (5 mg per kg); N12, after 12 days naloxone 20 mg per kg per day; M1, 4 h after first injection of morphine 8 mg per kg; M2, 4 h after injection of morphine 16 mg per kg on day 2; M3, after 16 h morphine (60 mg per kg per day) on day 3; M12, after continuation of morphine 60 mg per kg per day for 12 days; W2–W24, 2–24 hours after naloxone-precipitated opiate withdrawal; HS, 4 h after i.p. injection of 1.5 M saline, 1.8 ml per 100 g body weight. All experiments represent a minimum of five animals. *, $P < 0.05$ (Dunnett's multiple comparison test). Tissue sections were apposed to Kodak X-Omat AR film for 11 days and optical density measured using an image analysis system (Loats Associates, Westminster, Maryland). Optical densities were converted to copies of probe hybridized per μm^3 tissue using ^{35}S -labelled standards as previously described²⁸. The total number of copies of probe hybridized per paraventricular nucleus per slide was then calculated as the product of the area of the nucleus, the section thickness (12 μm) and the number of copies of probe per μm^3 .

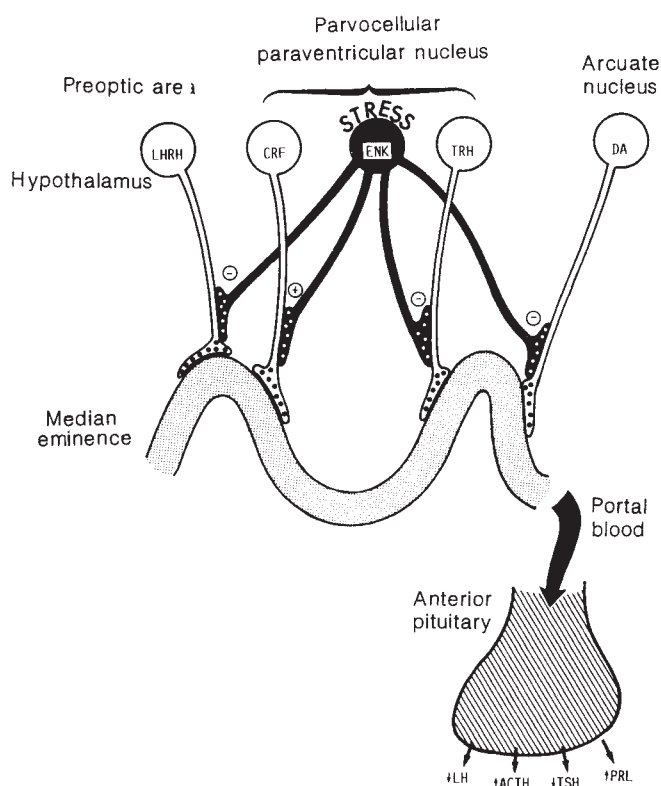


Fig. 3 Schematic representation of the possible contribution of paraventricular enkephalin neurons to the neuroendocrine response to stress in the rat. ACTH, adrenocorticotrophic hormone; CRF, corticotropin releasing factor; DA, dopamine; LH, luteinizing hormone; LHRH, luteinizing hormone releasing hormone; PRL, prolactin; TRH, thyrotropin releasing hormone; TSH, thyrotropin. Because CRF and enkephalin coexist in at least some of the parvocellular paraventricular cells⁹, they are probably coreleased in response to stress.

adrenergic, GABAergic or other neurons which project to these hypophysiotropic neurons.

There is now considerable evidence that endogenous opioid peptides are important in the stimulation of secretion of ACTH^{3,15}, and prolactin^{4,16-18}, and the inhibition of LH^{19,20} and TSH^{7,8} secretion. Their function in growth hormone secretion is less certain^{4,21-23}. It is particularly interesting that enkephalin can increase CRF release³, and opiate antagonists can block stress-induced stimulation of prolactin^{16,17} and inhibition of LH²⁰. Paraventricular enkephalin neurons may also modulate the posterior pituitary response to stress by acting on the dendrites of the adjacent magnocellular vasopressin and oxytocin cells which are found in the parvocellular zone of the paraventricular nucleus²⁴. The opiate antagonist naloxone has been shown to potentiate both the vasopressin and oxytocin response to stress²⁵. The rapid activation of hypothalamic enkephalin mRNA may therefore be an important contributory factor in the complex neuroendocrine response to stress.

Received 27 April 1987; accepted 12 June 1987.

1. Atweh, J. F. & Kuhar, M. J. *Brain Res.* **129**, 1-2 (1977).
2. Manson, P. L. *Prog. Brain Res.* **39**, 361-372 (1973).
3. Buckingham, J. C. & Cooper, T. A. *Neuroendocrinology* **38**, 411-417 (1984).
4. Bruni, J. F., Van Vugt, D. A., Marshall, S. & Meites, J. *Life Sci.* **21**, 461-466 (1977).
5. Kalra, S. P. *Endocrinology* **109**, 1805-1810 (1981).
6. Shaar, C. J., Frederickson, R. C. A., Dininger, N. B. & Jackson, L. *Life Sci.* **21**, 853-860 (1977).
7. Lomax, P. & George, R. *Brain Res.* **2**, 361-367 (1966).
8. Muraki, T., Nakadate, T., Tokunada, Y. & Kato, R. *J. Endocr.* **86**, 357-362 (1980).
9. Hökfelt, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 895-898 (1983).
10. Khachaturian, H., Lewis, M. E. & Watson, S. J. *J. comp. Neurol.* **220**, 310-320 (1983).
11. Merchenthaler, I., Maderdrut, J. L., Altschaler, R. A. & Petras, P. *Neuroscience* **17**, 325-348 (1986).
12. Fallon, J. H. & Leslie, F. M. *J. comp. Neurol.* **249**, 293-336 (1986).
13. Hisano, S., Daikoku, S., Yanaihara, N. & Shibasaki, T. *Brain Res.* **370**, 321-326 (1986).
14. Castanas, E. *et al. Brain Res.* **312**, 1-6 (1984).
15. Lotti, V. U., Kokka, N. & George, R. *Neuroendocrinology* **4**, 326-332 (1969).
16. Grandison, L. & Guidotti, A. *Nature* **270**, 357-359 (1977).
17. Rossier, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 666-669 (1980).
18. Gudelsky, G. A. & Porter, J. C. *Life Sci.* **25**, 1697-1702 (1979).
19. Rotsztein, W. H., Drouva, S. V., Pattou, E. & Kordon, C. *Nature* **274**, 281-282 (1978).
20. Brisky, K. P., Quigley, K. & Meites, J. *Life Sci.* **34**, 2485-2493 (1984).
21. Brisky, K. P., Quigley, K. & Meites, J. *Proc. Soc. exp. Biol. Med.* **177**, 137-142 (1984).
22. Drouva, S. V., Epelbaum, J., Tapia-Arancibia, L., Laplante, F. & Kordon, C. *Neuroendocrinology* **32**, 163-167 (1981).
23. Miki, N., Ono, M. & Shizume, K. *Endocrinology* **114**, 1950-1952 (1984).
24. Van den Pol, A. J. *comp. Neurol.* **206**, 317-345 (1982).
25. Carter, D. A., Williams, T. D. M. & Lightman, S. L. *J. Endocr.* **111**, 239-244 (1986).
26. Young, W. S. III, Bonner, T. I. & Brann, M. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9827-9831 (1986).
27. Yoshikawa, K., Williams, C. & Sabol, S. L. *J. biol. Chem.* **259**, 14301-14308 (1984).
28. Young, W. S. III, Mezey, E. & Siegel, R. E. *Neurosci. Lett.* **70**, 198-203 (1986).

Interaction between colour and motion in human vision

V. S. Ramachandran

Department of Psychology, C-009, University of California, San Diego, La Jolla, California 92093, USA

There is a wealth of anatomical¹⁻⁴ and psychological⁵⁻⁹ evidence which suggests that when people look at an object in the visual world, its various attributes such as colour, 'form', motion and depth are analysed by separate channels in the visual system. If so, how are these attributes put back together again to create a unified picture of the object? And if the object moves rapidly, how is perfect perceptual synchrony maintained between different features on its surface, if it is indeed true that they are being processed separately? Our evidence suggests that the visual system extracts certain conspicuous image features based on luminance contrast, and that the signals derived from these are then attributed to other features on the object, a process that we call 'capture'. Specifically, we find that when either illusory contours or random-dot patterns are moved in the vicinity of a colour-border, the colour border will also seem to move in the same direction even though it is physically stationary.

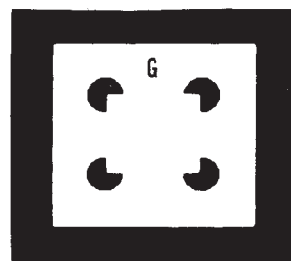


Fig. 1 An illusory square. Pie-shaped wedges are removed from four black disks and the disks are aligned appropriately to create the impression of a white square occluding four black disks in the background. G, green background.

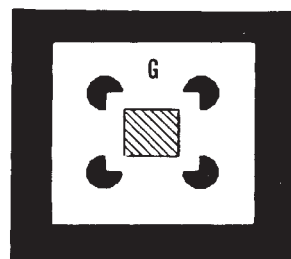


Fig. 2 The illusory square in this display is concentric with the smaller inner red square (cross-hatched). The red square is made equiluminous with the background and the illusory square is made to oscillate horizontally. Subjects report seeing the inner red square move in the same direction even though it is physically stationary. The illusion disappears if a small black spot is added to the centre of the red square. G, green background.

Illusory contours¹⁰⁻¹² can be produced by appropriately aligned black disks from which pie-shaped wedges have been removed (Fig. 1). The brain interprets this parsimoniously as a white square occluding four black disks in the background and not as four sectorized disks which have been aligned by the experimenter. It produces the impression of faint line joining the cut sectors even though no line exists physically, hence the name illusory contours.

Our stimuli were generated on an red-green-blue (RGB) monitor using an Amiga microcomputer. First, we generated a red square (subtending 0.6°) on a green background and took care to match the luminance of red and green regions using a colour-flow filter. A transparent overlay of the illusory square (Fig. 1) was then superimposed over the cathode ray tube (CRT) screen. The illusory figure subtended 0.75° so that its borders fell well outside the borders of the red square (Fig. 2). When we moved the illusory figure up and down or horizontally we found that the red square appeared to move with it in the same direction even though it was physically stationary. This effect (motion capture) was especially pronounced on eccentric fixation; when we fixated about 2.5° from the centre the effect could be seen vividly even for excursions of up to 0.5°.

A formal experiment along these lines was done on 11 naive subjects who had normal colour vision and were unaware of the purpose of the experiment. The entire display (including sectorized disks) was generated directly on the CRT using an Amiga microcomputer. The red square and the illusory square subtended 0.6° and 0.75° respectively and subjects fixated 2.5° away from the centre of the display. The luminance of the background green was held constant at 0.87 cd m⁻² (luminance was measured by using a United Detector Technology Model 161 photometer). The disks were black and their luminance was 0.03 cd m⁻². Each trial began with the disks and the inner square

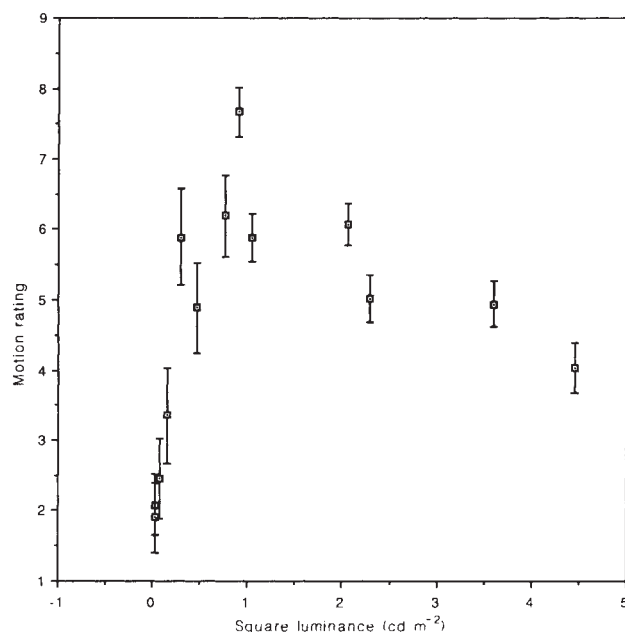


Fig. 3 Shows that optimum capture occurs when the red square is approximately equiluminous with the surround. Note however that some illusory movement was seen even when luminance-contrast was present (especially when the red square was brighter). Each datum point is based on 800 readings (11 subjects). Background luminance, 0.87 cd m^{-2} .

appearing simultaneously on the screen. The luminance value of the square was chosen randomly on any given trial. The outer disks then oscillated horizontally, making 0.3° excursions at a velocity of 0.2° per second. The subjects' task on each trial was to indicate the strength of perceived motion of the red square on a subjective scale of 0 to 9; with 0 being no capture and 9 being optimum capture. In making this judgment they were asked to rate the motion of inner red square alone. (Subjects were given a few practice trials before the experiment began.) Optimum capture was seen when the red square was approximately equiluminous with the background (Fig. 3). On the other hand, if luminance contrast was introduced across the border of the square, capture was reduced. This was especially true when the square was made darker.

We tried to determine whether the point at which optimum capture was observed did in fact correspond to equiluminance as defined by three other criteria. First, we generated a random-dot stereogram¹³ using the same luminance value for the green pixels and varied the luminance of red pixels gradually. Stereopsis disappeared¹⁴ when the red pixels had approximately the same luminance value (about 0.9 cd m^{-2}) as when optimum capture occurred. Similarly, we used a random-dot kinematogram and found a reduction of perceived motion at the same setting^{6,15}. And lastly, we used flicker-photometry to satisfy ourselves that this value did in fact correspond, at least roughly, to equiluminance as defined by more conventional criteria.

In our second experiment we studied the effects of varying the contrast of the disks while keeping the inner red square constantly equiluminous with the green background (background luminance in this experiment was 2.68 cd m^{-2}). The display was identical to Fig. 2 except that the disks were also red rather than black. Five of the subjects who participated in the previous experiment and four new subjects were asked to adjust the luminance of the disks until they saw minimum movement of the inner red square: that is, minimum capture.

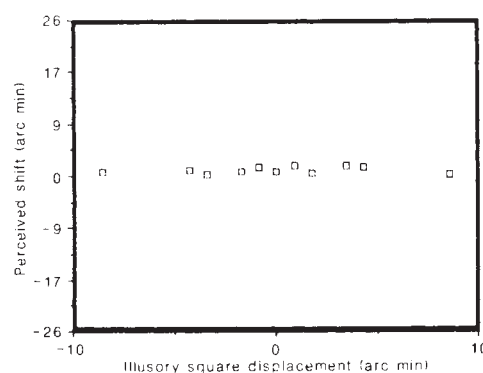


Fig. 4 Plot of the actual positional misalignment between the inner red square and sectorized disks against the perceived displacement of the inner square. Positive values indicate that the illusory square is shifted to the right. □, Mean of 832 readings (64 trials on each of 13 subjects). Error bars (not shown) were smaller than the square symbols. Results show that, at least for the spatial range over which motion capture was observed, there is no positional capture effect. Such an effect would be more likely to occur if the inner square were larger so that its borders were close to those of the illusory square.

Subjects could confidently make this adjustment and the mean luminance setting for 288 trials was 2.51 cd m^{-2} (s.d. = 0.37), which suggests that minimum capture occurs when the disks themselves are equiluminous with the background. We conclude that luminance derived motion signals can capture stationary chromatic borders but that moving chromatic borders cannot generate capture. Furthermore, three of our subjects spontaneously noted that even the disks themselves could no longer be seen to move at equiluminance which is consistent with the suggestion⁶ that colour borders provide only a weak input to human motion perception.

Is there a stationary analogue of motion capture? To find out, we used an illusory square that was stationary but displaced by varying amounts in relation to the central red square so that the two motifs were not exactly concentric. Would this spatial misalignment cause a shift in the perceived location of the red target? Thirteen subjects were asked to fixate 2.5° away from the centre of the display. On each trial a random spatial misalignment was introduced between the inner square and the sectorized disks. (All luminance values were identical to those in the previous experiment except that the disks were always black so that their luminance was not varied.) Two additional red squares were presented in the display, one above and one below the central red square, and the subject's task was to move these reference squares to set their position horizontally so that they were exactly aligned with the central square. (Each reference square was at least 2° above or below the sectorized disk array.) Figure 4 shows the results. Note that there is no positional capture effect at least for the spatial range over which motion capture was observed. Hence the motion effect was probably not a consequence of positional misalignment.

Another observation also supports this interpretation. We found that capture can also be produced if a sparse random dot pattern was used instead of an illusory square. A small red square was generated on a green background and the luminance of red and green were carefully matched using a colour-flow filter. A much larger pattern of sparse randomly positioned small black dots was then superimposed optically on the entire display and moved up and down by 0.5 – 1° . An opaque stationary window was used to occlude the margins of the random-dot pattern. As in our first experiment (Fig. 2), the red square was observed to move vividly in the same direction as the random

dot pattern even though it was physically stationary. The effect was reduced considerably if a small luminance contrast was introduced across the red-green border. Because a random-dot pattern has no continuous contours this effect cannot be ascribed to positional misalignment. Also, the effect demonstrates that any unambiguously moving surface can generate capture—not just illusory contours. The general rule seems to be that luminance derived motion signals capture chromatic borders in the vicinity.

These results establish capture as a robust perceptual phenomenon but we were concerned that tracking eye movements may have contributed to the effect. To rule out this possibility we generated two displays similar to Fig. 2 (using a mirror) so that the squares moved in opposite directions. Because the illusion could be seen in both displays simultaneously, we conclude that eye movements cannot explain the effect.

If tangential sections of area 18 in the macaque are stained for the cellular enzyme cytochrome oxidase, a distinct pattern of alternating broad and narrow stripes separated by non-staining interstripes is observed¹. Cells in the narrow stripes are sensitive to colour-contrast but not to orientation or direction of movement. Cells in the broad stripes are sensitive mainly to binocular disparity¹ and (probably) movement² of luminance borders, but are insensitive to colour borders. Perhaps when an object is suddenly displaced in the visual field the motion of its outline is picked up by the broad stripes and then spontaneously attributed to colour borders that happen to excite the adjacent narrow stripes. This would explain why colour borders are ineffective in generating capture. Similarly, capture of texture elements by moving sine-wave gratings¹⁶ and illusory contours¹⁷ can be explained by assuming that the motion signal from the broad bands are also attributed to stationary textures that are picked up by the adjacent interstripes in area 18. Cells in area 18 are known to respond to illusory contours¹⁸ but its not yet known whether these cells occur predominantly in the broad stripes.

Thus, when confronted with a moving target the visual system avoids redundancy by using the magnocellular luminance signals alone to compute motion and depth and then simply 'painting' the scene in an impressionistic way using the appropriate colours. This strategy allows economy of information processing and also ensures perceptual synchronization of form and colour as the object moves.

I thank R. M. Boynton, D. I. A. MacLeod, D. Hubel, M. Livingstone and F. H. C. Crick for discussions, J. Yellott and E. B. Ebbesen for facilities, and A. Diem for assistance. V.S.R. was supported by Academic Senate and Biomedical research funds from the University of California.

Received 2 March; accepted 2 June 1987.

- Hubel, D. H. & Livingstone, M. S. *Nature* **315**, 325–327 (1985).
- Van Essen, D. in *Cerebral Cortex* Vol. 3 (eds Peters, A. & Jones, E. G.) 259–329 (Plenum, New York, 1985).
- Allman, J. *Perception* **14**, 105–127 (1985).
- Shipp, S. & Zeki, S. M. *Nature* **315**, 322–324 (1985).
- Triesman, A. in *Physical and Biological Processing of Images* (eds Braddick, O. J. & Sleigh, A.) (Springer, Berlin, 1983).
- Ramachandran, V. S. & Gregory, R. L. *Nature* **275**, 55–56 (1978).
- Anstis, S. M. in *Handbook of Psychobiology* (eds Gazziniga, M. S. & Blakemore, C.) 269–317 (Academic, New York, 1975).
- Levinson, E. & Sekuler, R. *Vision Res.* **20**, 103–108 (1980).
- Held, R. & Shattuck, S. R. *Science* **174**, 314–316 (1971).
- Schumann, F. *Psychol. Stud. Lpz.* **1**, 1–32.
- Kanizsa, G. *Sci. Am.* **234**, 48–52 (1976).
- Gregory, R. L. *Nature* **238**, 51–52 (1972).
- Julesz, B. *Foundations of Cyclopean Perception* (University of Chicago Press, 1971).
- Lu, C. & Fender, D. H. *Invest. Ophthalmol. Vis. Sci.* **11**, 482–489 (1972).
- Cavanagh, P., Boeglin, J. & Favreau, O. *Perception* **14**, 151–162 (1985).
- Ramachandran, V. S. & Anstis, S. M. *Sci. Am.* **254**, 102–109 (1986).
- Ramachandran, V. S. *Perception* **14**, 127–134 (1985).
- Von Der Heydt, R., Peterhans, E. & Baumgartner, G. *Science* **224**, 1260–1262 (1984).

Parallel processing of motion and colour information

Thom Carney*, Michael Shadlen† & Eugene Switkes‡

* School of Optometry, Neurobiology Group, University of California, Berkeley, California 94720, USA

† Box G, Brown University, Providence, Rhode Island 02912, USA

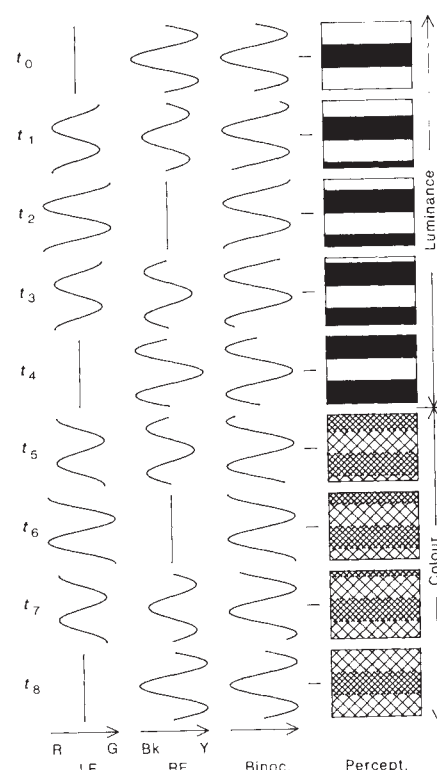
‡ Departments of Chemistry and Psychobiology, University of California, Santa Cruz, California 95064, USA

When the two eyes are confronted with sufficiently different versions of the visual environment, one or the other eye dominates perception in alternation. A similar situation may be created in the laboratory by presenting images to the left and right eyes which differ in orientation or colour. Although perception is dominated by one eye during rivalry, there are a number of instances in which visual processes nevertheless continue to integrate information from the suppressed eye^{1–6}. For example the interocular transfer of the motion after-effect is undiminished when induced during binocular rivalry⁶. Thus motion information processing may occur in parallel with the rivalry process. Here we describe a novel example in which the visual system simultaneously exhibits binocular rivalry and vision that integrates signals from both eyes. This apparent contradiction is resolved by postulating parallel visual processes devoted to the analyses of colour and motion information. Counterphased gratings are viewed dichoptically such that for one eye the grating is composed of alternating yellow and black stripes (luminance) while for the other it is composed of alternating red and green stripes (chrominance). When the gratings are fused, a moving grating is perceived. A consistent direction of motion can only be achieved if left and right monocular signals are integrated by the nervous system. Yet the apparent colour of the binocular percept alternates between red-green and yellow-black. These observations demonstrate the segregation of processing by the early motion system from that affording the perception of colour. Although, in this stimulus, colour information in itself can play no part in the cyclopean perception of motion direction, colour is carried along perceptually (filled in) by the moving pattern which is integrated from both eyes.

A pattern or texture defined solely by differences in colour, but not luminance, gives rise to a very weak perception of motion^{7–9}. For example, a moving isoluminant sinewave grating, seen as alternating red and green stripes, may appear virtually stationary with appropriately chosen red and green intensities¹⁰. On the basis of these and other observations¹¹ it can be argued that the early motion system processes spatial and temporal variations in pattern luminance with minimal use of colour information. Perception of a cyclopean stimulus requires the integration of information from both eyes but the occurrence of rivalry may preclude some aspects of such integration. By using a display that stimulates binocular rivalry as well as cyclopean motion perception, the separability of motion and colour information processing has been investigated.

A cyclopean motion display may be constructed by dichoptic presentation of counterphased sinewave gratings that differ by 90° in spatial and temporal phase between the two eyes. The luminance profiles of nine successive frames of two such gratings are depicted in the first two columns of Fig. 1. Simple algebraic summation of such patterns yields a drifting sinewave grating (column 3) which is precisely what observers perceive when these patterns are presented dichoptically¹². If the early motion system can process luminance contours without regard to chromaticity, cyclopean motion perception should be possible when counterphase gratings of different colours are viewed by the left and right eyes. Suppose one of the gratings consists of red and green stripes. Such a pattern may be constructed by combining red-black and green-black sinusoids in spatial anti-

Fig. 1 Nine successive frames (t_0 – t_8) of a cyclopean motion stimulus. **LE:** (left eye), luminance profiles, covering one and one-half spatial cycles, of the sinusoidally counterphased RG grating presented to the left eye. Arrow, direction of increasing luminance. The location of the perceptual labels R and G indicates that the green stripes are of higher intensity than the red stripes. Therefore the area of high luminance coincides with the green stripes and the area of low luminance coincides with the red stripes. **RE:** (right eye), luminance profiles of the sinusoidally counterphased yellow-black luminance grating presented to the right eye. Notice that the spatial patterns are shifted by one quarter cycle or 90° relative to the pattern viewed by the left eye. The temporal modulation is also shifted by one quarter cycle relative to the pattern viewed by the left eye, as evident from the flat luminance profile at t_0 in the left eye and at t_3 in the right eye. Arrow, direction of increasing luminance, which is also reflected in the position of the perceptual labels Bk for black and Y for yellow. **Binoc:** luminance profiles that are the sum of the corresponding left and right eye luminance profiles and represent the drifting binocular or cyclopean luminance grating. **Percept:** the boxes schematize the perception of a horizontal sinewave grating drifting upward. While the grating continually drifts upward, binocular rivalry is manifest as an alternation from the yellow-black appearance of the right eye grating to the red-green appearance of the left eye grating.



phase. If the contrasts of the red and green components are precisely balanced, relative to the human red and green sensitivities, an isoluminant red-green (RG) grating results. If the contrast of the red component is greater than that of the green, the resulting RG grating is equivalent to the sum of an isoluminant RG sinusoid and a luminance (yellow-black) sinusoid in phase with the red chromatic component. Conversely, a higher-contrast green sinusoid results in a luminance component which is in phase with the green chromatic component. Thus if one eye views a RG grating, it is only the luminance component of the RG grating, in combination with a luminance grating being presented to the other eye, that can determine the perceived direction of motion. Moreover, the 180° phase shift in the luminance component obtained by reversing the relative contrasts of the red and green components would be expected to reverse the perceived direction of motion (a somewhat analogous monocular stimulus has been described by Anstis and Cavanagh¹³). At the isoluminance point the perceived direction of motion is expected to be ambiguous or absent. As the R/G contrast ratio increases or decreases from the isoluminance point, the effective contrast of the luminance component will increase as should the subject's ability to consistently detect the direction of motion of the fused gratings.

Observers viewed horizontally oriented counterphased sinewave gratings through a modified mirror stereoscope. The luminance grating viewed by the right eye was of low contrast (10%) and shifted by 90° phase in space and time relative to the RG grating presented to the other eye. The ratio of the red and green component contrasts ranged from 0.9 to 1.6 for a total of 15 stimulus conditions. Each stimulus presentation lasted one second, after which the subject reported the perceived direction of motion. The results (Fig. 2) conform to the predictions, in that the perceived direction depends on the relative contrasts of the red and green components. Moreover this procedure followed by probit analysis¹⁴, provides a novel method for determining RG isoluminance for individual observers. Subjectively the perception is of a drifting RG grating whose velocity decreases as the isoluminance point is approached. These results indicate the binocular nature of the motion system and its dependence on luminance information.

Having a yellow-black grating in one eye and a red-green grating in the other eye ought to be ideal for eliciting binocular rivalry during prolonged viewing. To determine if rivalry would prevent binocular integration of luminance information and disrupt cyclopean motion perception, the gratings were presented continuously for 2 min while the subject indicated perceived grating colour and direction of motion. The direction of cyclopean motion was switched randomly by reversing the phase of the left or right gratings. To ensure smooth motion, the direction reversal was performed as counterphase gratings passed through zero contrast, so the event was not detected. The observer's task

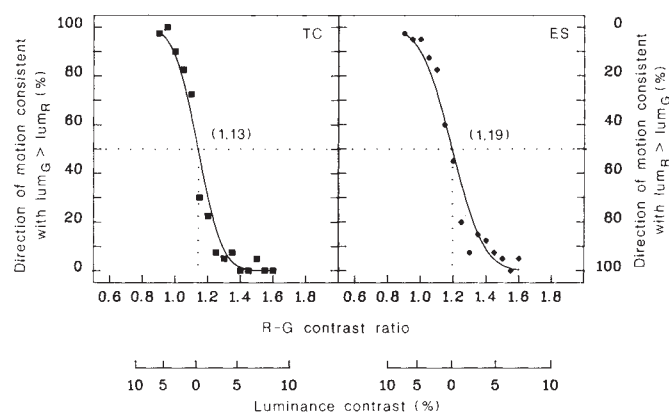
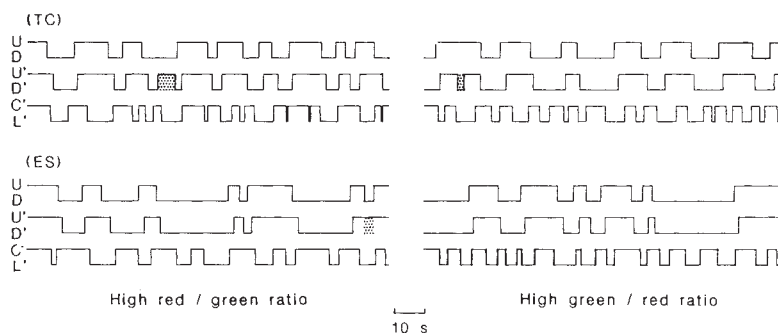


Fig. 2 The probability of perceiving the correct direction of motion based on the luminance components in the red and green stripes (lum_R and lum_G) as a function of the stimulus red-green contrast ratio. The stimuli viewed by the left and right eyes were counterphased sinewave gratings (1 cycle per degree and 2 cycles per second) in spatio-temporal quadrature (see Fig. 1). The stimuli were presented in random order for a total of 40 trials each. Probit analysis was used to fit a psychometric function to the data. The RG isoluminance point (ambiguous motion direction) for observers TC and ES was determined to be 1.13 and 1.19, respectively.

Fig. 3 Perceived colour and direction of motion during two 2-min periods for the same two observers. Left, data gathered in the presence of a red-green grating with the red sufficiently more intense than green to provide 10% luminance contrast. Right, data gathered in the presence of a red-green grating with green sufficiently more intense than red to provide a 10% luminance contrast. Top row: motion direction of the cyclopean luminance grating based on the interocular phase of the intense colour component and the luminance grating presented to the other eye. (U, up; D, down). Middle row: perceived direction of motion as indicated by button press. Except for a few short periods (stippled areas), perceived direction concurred with the direction of the cyclopean luminance grating (U', up; D', down). Bottom row: perceived colour of stimulus as indicated by button press (C', red-green; L', yellow-black).



was to signal the apparent direction (up or down) and the colour of the grating (red-green or yellow-black) by pressing the appropriate response box buttons. The apparent colour of the grating alternated between the colour of one and of the other monocular pattern and back again about every 8 s (Fig. 3). Often at the beginning of each alternation adjacent patches of yellow-black and red-green grating occurred briefly before complete rivalry. Rivalry does not interfere with direction of motion judgements requiring binocular integration. With the higher red contrast stimulus (Fig. 3, left panel), perceived direction is consistent with a luminance component in phase with the red component. The converse is true with a high green contrast stimulus (right panel).

Cyclopean motion perception demonstrates binocular integration of luminance information in the visual pathways mediating motion perception. The integrity of colour perception in the presence of conflicting information from the two eyes is manifested as rivalry. The result is consistent with the conjecture that luminance but not chrominance is the primary source of motion information. The inability of colour to convey motion information may also reflect an inability to convey absence of motion. Thus the flickering red-green and yellow-black patterns are carried along by the motion prescribed by their luminance components. In this sense colour is captured by luminance motion^{15,16}. These distinctions between colour and motion processes may reflect a neural organization in which colour and motion are processed separately. Recent anatomical and physiological studies suggest that a magnocellular(LGN)-striate-MT pathway may mediate many aspects of motion perception whereas colour and form are processed by a parvocellular(LGN)-striate-V2-V4 pathway¹⁷⁻²⁶. (Here LGN stands for lateral geniculate nucleus; MT for medial temporal area and V2 and V4 are prestriate visual areas.

We thank Dr Paradiso for criticisms of the manuscript. M.S. and T.C. were supported by an NIH fellowship. This research was also supported by an NIH grant to E. Switkes.

Received 29 January; accepted 22 May 1987.

1. Blake, R. & Fox, R. *Nature* **249**, 488-490 (1974).
2. White K. D., Petry, H. M., Riggs, L. A. & Miller, J. *Vision Res.* **18**, 1201-1215 (1978).
3. Wade, N. J. & Wenderoth, P. *Vision Res.* **18**, 827-835 (1978).
4. Blake, R. & Overton, R. *Perception* **8**, 143-152 (1979).
5. Lehmkuhle, S. W. & Fox, R. *Vision Res.* **15**, 855-859 (1975).
6. O'Shea, R. P. & Crassini, B. *Vision Res.* **21**, 801-804 (1981).
7. Ramachandran, V. S. & Gregory, R. L. *Nature* **275**, 55-56 (1978).
8. Cavanagh, P., Boeglin, J. & Favreau, O. E. *Perception* **14**, 151-162 (1985).
9. Cavanagh, P. & Favreau, O. E. *Vision Res.* **25**, 1595-1602 (1985).
10. Gregory, R. L. & Harris, J. P. *Nature* **307**, 729-730 (1984).
11. Cavanagh, P., Tyler, C. W. & Favreau, O. E. *J. opt. Soc. Am. A* **1**, 893-899 (1984).
12. Shadlen, M. & Carney, T. *Science* **232**, 95-97 (1986).
13. Anstis, S. & Cavanagh, P. in *Colour Vision: Physiology and Psychophysics* (eds Mollon, J. D. & Sharpe, L. T.) 155-166 (Academic, New York, 1983).
14. Finney, D. J. *Probit Analysis* (Cambridge University Press, 1971).
15. Ramachandran, V. S. *Percept. Psychophys.* **39**, 361-373 (1986).

16. Ramachandran, V. S. *Perception* **14**, 127-134 (1985).
17. Shipp, S. & Zeki, S. M. *Nature* **315**, 322-324 (1985).
18. Van Essen, D. C. in *Cerebral Cortex* (eds Petters, A. & Jones, E. G.) Vol. 3, 259-329 (Plenum, New York, 1985).
19. Hubel, D. H. & Livingston, M. S. *Nature* **315**, 325-327 (1985).
20. DeYoe, E. A. & Van Essen, D. C. *Nature* **317**, 58-61 (1985).
21. Burkhalter, A. & Van Essen, D. C. *J. Neurosci.* **6**, 2327-2351 (1986).
22. Albright, T. D. *J. Neurophysiol.* **52**, 1106-1130 (1984).
23. Felleman, D. J. & Kaas, J. H. *J. Neurophys.* **52**, 499-513 (1984).
24. Maunsell, J. H. R. & Van Essen, D. C. *J. Neurophysiol.* **49**, 1127-1147 (1983).
25. Weller, R. E., Wall, J. T. & Kaas, J. H. *J. comp. Neurol.* **228**, 81-104 (1984).
26. Zeki, S. M. *Nature* **274**, 423-428 (1978).

Site-specific phosphorylation induces disassembly of vimentin filaments *in vitro*

Masaki Inagaki*, Yoshimi Nishi†‡, Kimiko Nishizawa*, Mutsushi Matsuyama† & Chikako Sato*

* Laboratory of Experimental Radiology and
† Ultrastructure Research, Aichi Cancer Center Research Institute,
Chikusa-ku, Nagoya, Aichi 464, Japan
‡ Life Information Analysis Center, Komaki, Aichi 485, Japan

Intermediate filaments are a major component of the cytoskeleton of eukaryotic cells. Although there appear to be at least five distinct classes of these filaments, cells of mesenchymal origin and most cells in culture contain the intermediate filament composed of the subunit protein vimentin¹. Vimentin exists in a nonphosphorylated as well as in a phosphorylated form¹⁻⁴, and there is increased phosphorylation of this protein when the filament undergoes marked redistribution in various cells⁵⁻⁸. The role of phosphorylation on assembly-disassembly and organization of the vimentin filament has remained obscure. We report here a stable and purified system allowing biochemical examination of vimentin filament assembly and disassembly. Using this *in vitro* system, we carried out stoichiometrical phosphorylations, using purified protein kinases. We obtained evidence for site-specific, phosphorylation-dependent disassembly of the vimentin filament.

The cytoplasmic regulatory mechanisms that govern assembly-disassembly of vimentin filament are obscure. One possible mechanism is the post-translational modification (phosphorylation) of vimentin¹⁻⁸. As phosphorylation-dephosphorylation reactions are involved in various biological control phenomena^{9,10}, it is possible that the assembly-disassembly of vimentin into cellular structures might be regulated by a similar event. As a first step towards defining the regulatory mechanism, we carried out an analysis of phosphorylation of vimentin, using cAMP-dependent protein kinase, smooth muscle myosin light chain kinase, liver calmodulin-dependent glycogen synthase kinase, casein kinase I, casein kinase II and protein kinase C. Only cAMP-dependent protein kinase and protein kinase C

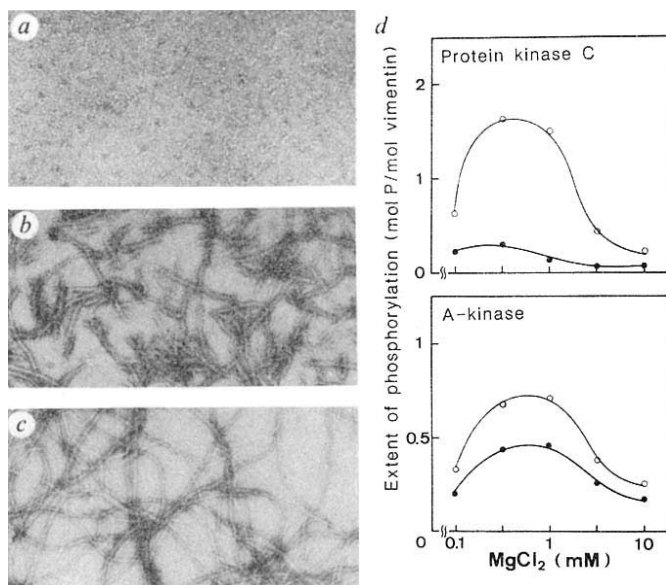


Fig. 1 Electron microscopy of monomeric vimentin incubated with 0.3 mM MgCl₂ (a) or 10 mM MgCl₂ (b) in the absence of kinase for 10 min; c, normal vimentin filament. All samples were negatively stained. Scale bar, 200 nm. d, Monomeric vimentin (○) and vimentin filaments (●) were phosphorylated using either protein kinase C (top diagram) or the catalytic subunit of cAMP-dependent protein kinase (bottom diagram) at various concentrations of MgCl₂.

Methods. a–c, The preparations were placed directly on carbon-film-coated specimen grids and negatively stained with 1% or 2% uranyl acetate or 20% sodium phosphotungstate (pH 7.2). d, Monomeric vimentin and vimentin filament were phosphorylated by protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase and the reactions were halted at 10 min. Crude vimentin preparations from Furth's murine mastocytoma cells were prepared essentially as described by Zackroff and Goldman¹⁹. Vimentin was purified from crude vimentin preparations by a modification of reported methods²⁰. Between 3 and 6 mg purified vimentin was prepared from 20 g mastocytoma cells. The catalytic subunit of cAMP-dependent protein kinase was prepared from bovine heart according to Beavo *et al.*²¹. Protein kinase C was prepared from rabbit brain by the method of Inagaki *et al.*²². Vimentin (5.4 μM) was phosphorylated by incubation with 5 μg ml⁻¹ protein kinase C, 0.1 mM [γ -³²P] ATP, 0.8 mM CaCl₂, 30 mM NaCl, 50 μg ml⁻¹ phosphatidylserine, 25 mM Tris-HCl pH 7.0 and various MgCl₂ concentrations at 25 °C, or by incubation with 5 μg ml⁻¹ of the catalytic subunit of cAMP-dependent protein kinase, 0.1 mM [γ -³²P] ATP, 30 mM NaCl, 25 mM Tris-HCl, pH 7.0 and various MgCl₂ concentrations at 25 °C.

phosphorylated vimentin at significant rates. Purified vimentin was obtained by extraction of the crude intermediate filament preparation from Furth's murine mastocytoma cells with 8 M urea and the subsequent chromatography on DEAE cellulose and CM cellulose columns run in the presence of urea. Removal of the urea by dialysis for 24–36 h produced a protein that was mostly monomeric in 10 mM Tris-HCl, pH 8.8. The monomeric vimentin (protein concentration of 0.3–1 mg ml⁻¹) then assembled into 7–12 nm diameter filaments during dialysis for 24–36 h against 150 mM NaCl, 10 mM imidazole-HCl, pH 7.0^{11,12}. In negatively stained preparations, the reconstituted vimentin filaments were mostly smooth-curved and were morphologically similar to native intermediate filaments.

The assembly and disassembly of vimentin and formation of 10 nm filaments from highly purified vimentin are influenced by ionic strength^{11,12}, pH¹² and MgCl₂ concentration in the buffer. To define the regulatory activity of phosphorylation itself for the assembly–disassembly of the vimentin filament, we examined the effects of ionic strength and MgCl₂ concentration on each

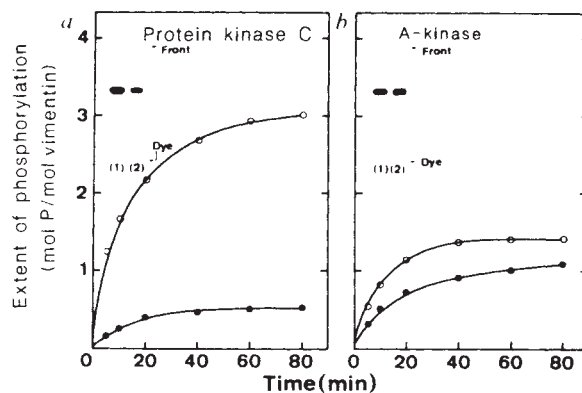


Fig. 2 Time course of phosphorylation of monomeric vimentin (○) and vimentin filament (●) by a, protein kinase C and b, the catalytic subunit of cAMP-dependent protein kinase. Inset, autoradiograms of SDS-polyacrylamide slab gels of monomeric vimentin (line 1) and vimentin filament (line 2) phosphorylated by either protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase samples taken at 30 min.

Methods. Monomeric vimentin or vimentin filaments were phosphorylated by protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase, as described in the legend for Fig. 1, with 0.3 mM MgCl₂. The sample was boiled in a water bath for 1 min and subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE) under the conditions described by Laemmli²³. The separating and stacking gels contained 12.5 and 4% acrylamide, respectively. The gel was stained with Coomassie Brilliant Blue, destained and exposed to a Kodak X-Omat RPX-ray film.

state of vimentin, at a physiological pH (25 mM Tris-HCl pH 7.0). The properties of the vimentin filament, visualized by electron microscopy, remained constant between 20 mM and 150 mM NaCl and the properties of monomeric vimentin also remained fairly constant between 0 mM and 50 mM NaCl, at low total MgCl₂ concentrations (<1 mM), when the samples were incubated for up to 1.5 h at 25 °C. MgCl₂ concentration appeared to have the most consistent effect on polymerization. Increasing the MgCl₂ concentration (>3 mM) favoured formation of the vimentin filament within 10 min. As shown in Fig. 1b the MgCl₂-induced vimentin filament is smooth-curved like the native 10 nm filaments but thicker, as ~18–25 nm in diameter. The dramatic change in the state of vimentin as a function of MgCl₂ concentration was reflected also by alteration in the rate of vimentin phosphorylation by the protein kinases (Fig. 1d). ATP (0.01–1 mM) alone had no effect on either state of vimentin. Therefore, all experiments related to vimentin phosphorylations were performed in a reaction mixture containing 25 mM Tris-HCl, pH 7.0, 0.1 mM ATP, 30 mM NaCl and 0.3 mM MgCl₂.

As shown by the samples taken at 30 min, protein kinase C incorporated 2.5 mol phosphate into 1 mol vimentin when monomeric vimentin was used as a substrate and 0.5 mol phosphate into 1 mol vimentin when the vimentin filament was used as a substrate. The catalytic subunit of cAMP-dependent protein kinase could incorporate 1.3 mol phosphate into 1 mol vimentin when monomeric vimentin was used as a substrate and 0.8 mol phosphate into 1 mol vimentin when the vimentin filament was used as a substrate (Fig. 2). The effects of these phosphorylations on the state of vimentin were determined by electron microscopy of negatively stained vimentin samples. Phosphorylation of monomeric vimentin by either protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase had little effect on the depolymerized configuration. In contrast, phosphorylation of the vimentin filament by the catalytic subunit of cAMP-dependent protein kinase, but not protein kinase C, dramatically induced a transition toward the depolymerized state (the monomeric state). Typical images of the phosphory-

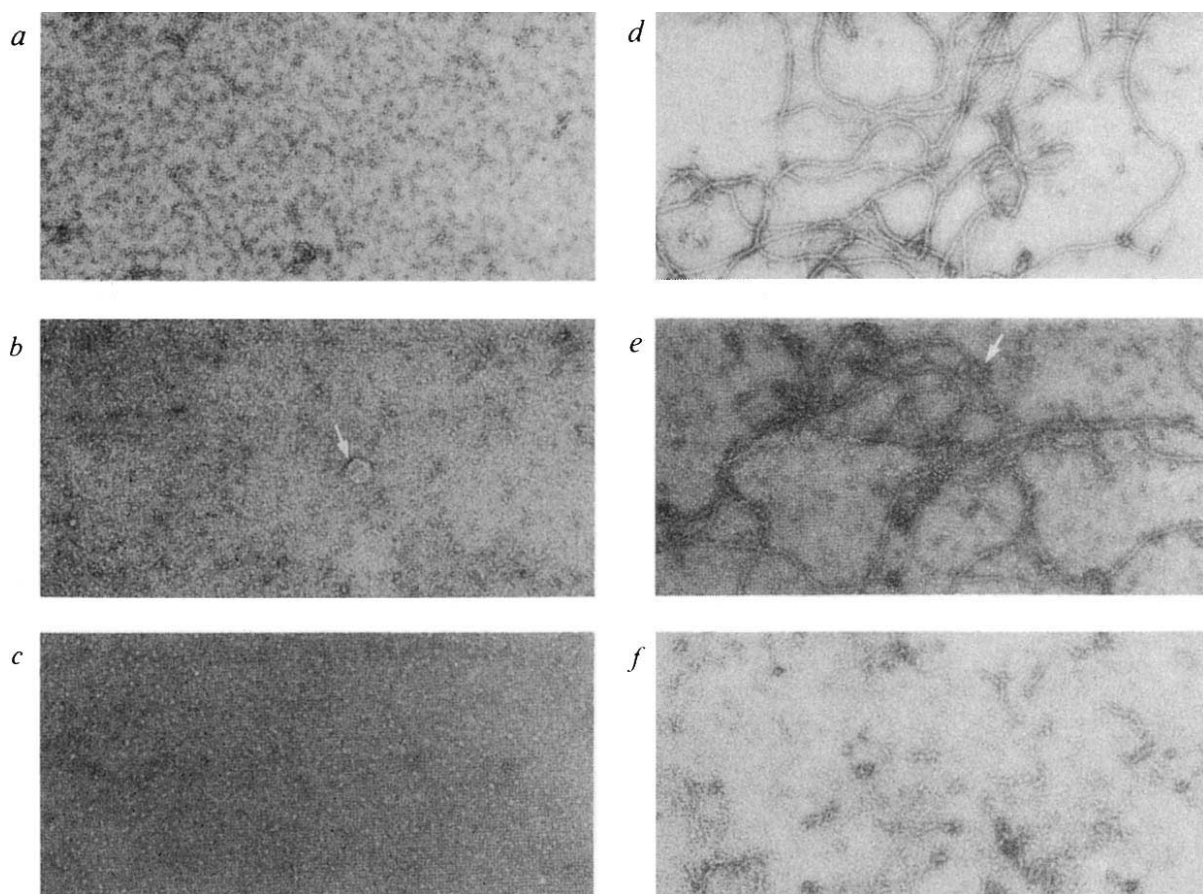


Fig. 3 Electron microscopy of the negatively stained vimentin samples phosphorylated by protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase. *a*, Unphosphorylated monomeric vimentin; *b*, monomeric vimentin phosphorylated by protein kinase C; *c*, monomeric vimentin phosphorylated by the catalytic subunit of cAMP-dependent protein kinase; *d*, unphosphorylated vimentin filament; *e*, vimentin filament phosphorylated by protein kinase C; *f*, vimentin filament phosphorylated by the catalytic subunit of cAMP-dependent protein kinase. Scale bars, 200 nm. Electron microscopy of negatively stained vimentin samples was performed as described in the legend for Fig. 1. Arrows in *b* and *e* indicate vesicles of phosphatidylserine. Vimentin was phosphorylated as described in the legend of Fig. 2. The samples taken at 30 min were negatively stained. In the control experiment, the same assay was carried out without ATP.

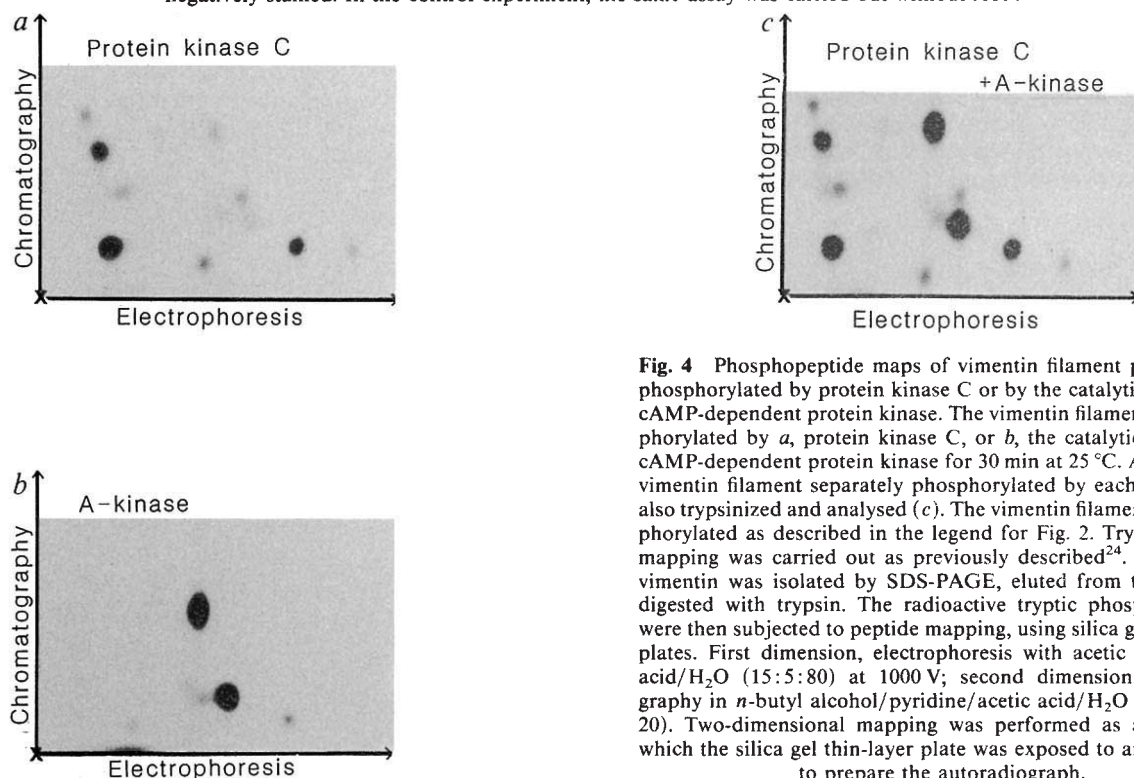


Fig. 4 Phosphopeptide maps of vimentin filament preparations phosphorylated by protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase. The vimentin filament was phosphorylated by *a*, protein kinase C, or *b*, the catalytic subunit of cAMP-dependent protein kinase for 30 min at 25 °C. A mixture of vimentin filament separately phosphorylated by each kinase was also trypsinized and analysed (*c*). The vimentin filament was phosphorylated as described in the legend for Fig. 2. Tryptic peptide mapping was carried out as previously described²⁴. Radioactive vimentin was isolated by SDS-PAGE, eluted from the gel, and digested with trypsin. The radioactive tryptic phosphopeptides were then subjected to peptide mapping, using silica gel thin-layer plates. First dimension, electrophoresis with acetic acid/formic acid/H₂O (15:5:80) at 1000 V; second dimension, chromatography in *n*-butyl alcohol/pyridine/acetic acid/H₂O (32.5:25:5:20). Two-dimensional mapping was performed as above, after which the silica gel thin-layer plate was exposed to an X-ray film to prepare the autoradiograph.

lated samples are shown in Fig. 3. As vimentin is very susceptible to proteolysis and limited proteolysis of vimentin results in deletion of the filament-assembly site of the protein^{13,14}, we used two-dimensional gels to rule out the possibility that the preparation of the catalytic subunit of cAMP-dependent protein kinase used contained a protease impurity. The phosphovimentin samples described above migrated at the same molecular weight as that of vimentin and were slightly more acidic than the vimentin. Proteolytic breakdown products of vimentin were never detected (data not shown). Neither the catalytic subunit of cAMP-dependent protein kinase nor MgATP separately produced disassembly of the vimentin filament. Complete disassembly occurred when vimentin was maximally phosphorylated by the catalytic subunit of the cAMP-dependent protein kinase with MgATP.

We examined whether the site on vimentin phosphorylated by the catalytic subunit of cAMP-dependent protein kinase is different from that phosphorylated by protein kinase C. After each vimentin preparation was maximally phosphorylated by either enzyme the vimentin samples were fractionated on SDS-polyacrylamide gels, eluted and digested with L-1-tosylamido-s-phenylethyl chloromethyl ketone-trypsin. The tryptic fragments were analysed by two-dimensional peptide mapping, using electrophoresis in the first dimension and thin layer chromatography in the second one. As shown in Fig. 4, different maps of phosphopeptides were obtained by tryptic hydrolysis from vimentin filament preparations phosphorylated by each kinase. The same maps of phosphopeptides were obtained from monomeric vimentin and vimentin filament preparations phosphorylated by protein kinase C as well as the catalytic subunit of cAMP-dependent protein kinase. Serine was the major phosphoamino acid resulting from the phosphorylation of vimentin by either protein kinase C or by the catalytic subunit of cAMP-dependent protein kinase (data not shown).

The cells of most, if not all, vertebrates contain three classes of cytoplasmic filaments: microtubules, microfilaments and intermediate filaments. Whereas a great deal is known of the biochemistry, structure, assembly-disassembly, and function of microtubules and microfilaments¹⁵⁻¹⁸, intermediate filaments remain the least understood of these protein fibrils. Our results indicate that reversible enzymatic phosphorylations of vimentin may be involved in modulating the state of polymerization of vimentin and its reversible disassembly in various cell functions.

We thank Dr Y. Nishizuka for encouragement and Drs H. Nakamura, K. Ohtsuka, S. Nakasu and M. Ohara for reading the manuscript and for helpful discussions. We also thank S. Tokumasu for secretarial services. This work was supported in part by grants from the Ministry of Education, Science and Culture, Japan.

Received 21 April; accepted 26 June 1987.

1. Lazarides, E. A. *Rev. Biochem.* **51**, 219-250 (1982).
2. O'Connor, C. M., Gard, D. L. & Lazarides, E. *Cell* **23**, 135-143 (1978).
3. Cabral, F. & Gottesman, M. M. *J. biol. Chem.* **254**, 6203-6206 (1979).
4. Gard, D. L., Bell, P. B. & Lazarides, E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3894-3898 (1979).
5. Evans, R. M. & Fink, L. M. *Cell* **29**, 43-52 (1982).
6. Celis, J. E., Mose Larsen, P., Fey, S. J. & Celis, A. J. *Cell Biol.* **97**, 1429-1434 (1983).
7. Gard, D. L. & Lazarides, E. *Molec. cell. Biol.* **2**, 1104-1114 (1982).
8. Evans, R. M. *J. biol. Chem.* **259**, 5372-5375 (1984).
9. Krebs, E. G. & Beavo, J. A. A. *Rev. Biochem.* **48**, 923-959 (1979).
10. Rubin, C. S. & Rosen, O. M. A. *Rev. Biochem.* **44**, 831-922 (1975).
11. Huiatt, T. W., Robson, R. M., Arakawa, N. & Stromer, M. H. *J. biol. Chem.* **255**, 6981-6989 (1980).
12. Ranner, W. et al. *J. molec. Biol.* **149**, 285-306 (1981).
13. Nelson, W. J. & Traub, P. *Molec. cell. Biol.* **3**, 1146-1156 (1983).
14. Kaufman, E., Weber, K. & Geisler, N. *J. molec. Biol.* **185**, 733-742 (1985).
15. Clarke, M. & Spudich, J. A. A. *Rev. Biochem.* **46**, 797-822 (1977).
16. Mohri, H. *Biochim. biophys. Acta* **456**, 85-127 (1976).
17. Olmsted, J. B. & Borisy, G. G. A. *Rev. Biochem.* **42**, 507-540 (1973).
18. Pollard, T. D. & Weihing, R. R. *Crit. Rev. Biochem.* **2**, 1-65 (1974).
19. Zackroff, R. V. & Goldman, R. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6226-6230 (1979).
20. Geisler, N. & Weber, K. *FEBS Lett.* **125**, 253-256 (1981).
21. Beavo, J. A., Bechtel, P. J. & Krebs, E. G. *Meth. Enzym.* **38**, 299-308 (1974).
22. Inagaki, M., Watanabe, M. & Hidaka, H. *J. biol. Chem.* **260**, 2922-2925 (1985).
23. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
24. Inagaki, M., Kawamoto, S. & Hidaka, H. *J. biol. Chem.* **259**, 14321-14323 (1984).

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the
Clearinghouse, please fill out and mail
back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about
UMI Article Clearinghouse. I am interested in
electronic ordering through the following
system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnType ☐ OCLC ILL Subsystem
☐ Other (please specify) _____
☐ I am interested in sending my order by mail.
☐ Please send me your current catalog and user instructions for
the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

Mail to: University Microfilms International
300 North Zeeb Road, Box 91 Ann Arbor, MI 48106